

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:15:20 ; Search time 9505 Seconds  
(without alignments)  
10825.501 Million cell updates/sec

Title: us-09-105-117k-1

Perfect score: 2374

Sequence: 1 ccatttgctgaaggtgttac.....ttcttcacaaaggatctct 2374

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hig.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pi.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_ats.\*  
12: gb\_av.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_ats.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	2374	100.0	2374	1	CGLYSEG	X96471	C-glutaminc
2	2374	100.0	2374	6	A93933	A93933	Sequence 2
3	2374	100.0	33150	1	AP005277	AP005277	Coryneb
4	2374	100.0	34980	6	AX127147	AX127147	Sequence
5	993	41.8	993	6	AX063767	AX063767	Sequence
6	993	41.8	993	6	AX244055	AX244055	Sequence
7	870	36.6	870	6	AX123540	AX123540	Sequence
8	870	36.6	870	6	BD165657	BD165657	Novel pol
9	822	34.6	822	6	AX063771	AX063771	Sequence
10	822	34.6	822	6	AX244059	AX244059	Sequence
11	818.6	34.5	308650	1	AP005218	AP005218	Coryneb
12	798.8	33.6	1771	1	AB083133	AB083133	Coryneb
13	724.4	30.5	349535	1	AX248357	AX248357	Coryneb
14	711	29.9	711	6	AX643028	AX643028	Sequence
15	708	29.8	708	6	AX123539	AX123539	Sequence
16	708	29.8	708	6	BD165656	BD165656	Novel pol
17	698.6	29.4	1568	6	ES4483	ES4483	Heat-resist
18	698.6	29.4	1568	6	AX643030	AX643030	Sequence
19	649	27.3	1083	6	AX813971	AX813971	Sequence
20	627	26.4	627	6	AX123538	AX123538	Sequence
21	627	26.4	627	6	BD165655	BD165655	Novel pol
22	485.2	20.4	993	6	AX067087	AX067087	Sequence
23	170	7.2	345783	1	AP003001	AP003001	Mesrhizo
24	164.4	6.9	8472	1	AE008076	AE008076	Agrobacte
25	164.4	6.9	13051	1	AE009111	AE009111	Agrobacte
26	159.4	6.7	302007	1	SC0939132	SC0939132	Streptomy
27	140.6	5.9	303550	1	SC0939131	SC0939131	Streptomy
28	138.6	5.7	306250	1	SME591786	SME591786	Sinorhizo
29	123.8	5.2	347356	1	EX640437	EX640437	Bordetell
30	121.8	5.2	348171	1	EX640412	EX640412	Bordetell
31	122.8	5.2	14844	1	MBU34849	MBU34849	Mycobacteri
32	122.8	5.2	15239	6	I86263	I86263	Sequence 17
33	122.8	5.2	306050	1	EX248341	EX248341	Mycobacte
34	122.2	5.1	348251	1	EX640423	EX640423	Bordetell
35	121.2	5.1	14869	1	AE007056	AE007056	Mycobacte
36	121.2	5.1	346186	1	EX842578	EX842578	Mycobacte
37	120.8	5.1	298300	1	AP005025	AP005025	Streptomy
38	116.4	4.9	11120	1	RFA311775	RFA311775	Rhodococc
39	115.2	4.9	298300	1	AP005025	AP005025	Streptomy
40	113	4.8	303550	1	SC0939131	SC0939131	Streptomy
41	112.6	4.7	10871	1	AE011965	AE011965	Xanthomon
42	112.4	4.7	10336	1	AE005530	AE005530	Escherich
43	112.2	4.7	12070	1	AB011413	AB011413	Streptomy
44	111.2	4.7	10278	1	AE015317	AE015317	Shigella
45	111.2	4.7	266658	1	AP002563	AP002563	Escherich

# ALIGNMENTS

RESULT 1	CGLYSEG	2374 bp	DNA	linear	BCT 19-MAR-2001
LOCUS	C-glutaminc	lysE and lysG genes.			
DEFINITION	X96471				
ACCESSION	X96471.1	GI:1729753			
VERSION	lysE gene; lysG gene; Lysine export regulator protein; lysine				
KEYWORDS	exporter protein; Lysine governor.				
SOURCE	Corynebacterium glutamicum				
ORGANISM	Corynebacterium glutamicum				
REFERENCE	Bacteria; Actinobacteria; Actinomycetales;				
AUTHORS	Corynebacterineae; Corynebacteriaceae; Corynebacterium.				
TITLE	Vrljic.M., Sahm.H. and Eggeling.L.				
	A new type of transporter with a new type of cellular function:				

JOURNAL L-lysine export from *Corynebacterium glutamicum*  
 MEDLINE Mol. Microbiol. 22 (5), 815-826 (1996)  
 PUBMED 97126810  
 REFERENCE 8971704  
 AUTHORS 2 (bases 1 to 2374)  
 DIRECT Vrjic, M.M.  
 SUBMISSION  
 TITLE Submitted (07-MAR-1996) M.M. Vrjic, Institut fuer Biotechnologie  
 JOURNAL 1, Forschungszentrum Juelich, Postfach 1913, D-52425 Juelich, FRG  
 FEATURES Location/Qualifiers  
 source 1..2374

/organism="Corynebacterium glutamicum"

/mol\_type="genomic DNA"

/strain="R127"

/db\_xref="taxon:1718"

complement (82..954)

/gene="lysG"

complement (82..954)

/gene="lysG"

/note="lysine governor"

/codon\_start=1

/transl\_table=11

/product="lysine export regulator protein"

/protein\_id="CAA65323.1"

/db\_xref="GI:1729754"

/db\_xref="GOA:P94632"

/db\_xref="SWISS-PROT:P94632"

/translation="MNPQLDTLISLIDRSGFEGASLALSISPSAVSRVKALEHRVG  
 RVLSRTQAKATEAGEVIVQAKRMVLQAKTQKLSGLRAEIPLATINADLSLSTW  
 FPPVNEVASGGATITLRLDEAHTLSLRGQVGLAVTRANFVAGREVLEGTMR  
 HLAITPSLRDYMVDGLDWAAPVLRGPKVQLQDROLDRGVDPGVRVRVSVPS  
 ALRGEALRLGLGWLPEPQAAAPLKAQEVILLIDPDTMYWQRVLESRLARL  
 TDVVDAAIEGLRP"

1025..1726

/gene="lysG"

1025..1726

/gene="lysG"

/codon\_start=1

/transl\_table=11

/product="lysine exporter protein"

/protein\_id="CAA65324.2"

/db\_xref="GI:13397387"

/db\_xref="GOA:P94633"

/db\_xref="SWISS-PROT:P94633"

/translation="MEIFITGLIGASLISIGSPNVLVIKQIRGLIAVLVCLL  
 SDVFLFAGTIGVDLSNAEPIVLDIMRGGLAYLFWAKADRMKNVRFQIE  
 ETSPTVPDPTPLGSAAVATDTRNRVRVSVSDKQKRVKPMALVITVFNPNAYLDA  
 FVFIGVGAQYGGDGRWIPFAAGAFASLIWFPLVGFAGAAALSRPLSSPKVWRWVNV  
 AVYMTALAIKMLMG"

complement (11723..2352)

/gene="orf3"

complement (11723..2352)

/gene="orf3"

/codon\_start=1

/transl\_table=11

/protein\_id="CAA65325.1"

/db\_xref="GI:1729756"

/db\_xref="SWISS-PROT:P94634"

/translation="MYALRDIVAGSKALVYGISSGPELPABAEFWAEGCPPLIHQ  
 PYSIINRWPEPDDGNLQSAANGLGVTFAPLQGLLTDKYLDGTPGSRASQ  
 GKSLSGLMNNYNDVRYKLDIAQERGOSLAQWALVLEQGYGADTVTSALIGA  
 SSVEQLDNLSDLSNLFSDAELEIDESHDAGINIWAKATDSKTREN"

# ORIGIN

Query Match 100.0%; Score 2374; DB 1; Length 2374;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATTTCGTGAGGTGTTACTCTGCTGCGCCCAATTCCTGCGGCGGAGAGTGTGAAAC 60

DB 1 CCATTTCGTGAGGTGTTACTCTGCTGCGCCCAATTCCTGCGGCGGAGAGTGTGAAAC 60

QY 61 CCTGAACCTTTTCAGAACTAACTAAGGCGCGCAATCCCTCGATTGCTGCAACACGCGC 120

DB 61 CCTGAACCTTTTCAGAACTAACTAAGGCGCGCAATCCCTCGATTGCTGCAACACGCGC 120  
 QY 121 GTCTGTGAGTCTAGCTAGAGATCTAGATTTCCAGGCGGCATCTGTCGCAATACATCGGTG 180  
 DB 121 GTCTGTGAGTCTAGCTAGAGATCTAGATTTCCAGGCGGCATCTGTCGCAATACATCGGTG 180  
 QY 181 GTCAATGGGTATCTCATCGAGGAGATCACTTCTCTGCTTTTAGCATGGAGGAGCTTG 240  
 DB 181 GTCAATGGGTATCTCATCGAGGAGATCACTTCTCTGCTTTTAGCATGGAGGAGCTTG 240  
 QY 241 GGTTCGGGAAGAAGTCCCAACCAAGGCGCTCGCGCAATTCGCTCACAAAACCTTTCGCG 300  
 DB 241 GGTTCGGGAAGAAGTCCCAACCAAGGCGCTCGCGCAATTCGCTCACAAAACCTTTCGCG 300  
 QY 301 CGACGGGCAATGGATAGCGCTCTGCGCCCAACAGGACCATCGACGCGCCGCTCCAGGTG 360  
 DB 301 CGACGGGCAATGGATAGCGCTCTGCGCCCAACAGGACCATCGACGCGCCGCTCCAGGTG 360  
 QY 361 AGGCTCTGAGCACATCTTTGGGACCGGAGCGTAAAGCGGGCATCGCAGCCCATCTAG 420  
 DB 361 AGGCTCTGAGCACATCTTTGGGACCGGAGCGTAAAGCGGGCATCGCAGCCCATCTAG 420  
 QY 421 TTTCCCATCAACATGATAGGCATCCCGCAATGAGGGGTTGCAATGCGCAAGTGGCGAT 480  
 DB 421 TTTCCCATCAACATGATAGGCATCCCGCAATGAGGGGTTGCAATGCGCAAGTGGCGAT 480  
 QY 481 GGTTCGAAGTTCTACTTACTTCCATCCCGCACCGGATTAGCTTCAGGGTTACCGGCTCC 540  
 DB 481 GGTTCGAAGTTCTACTTACTTCCATCCCGCACCGGATTAGCTTCAGGGTTACCGGCTCC 540  
 QY 541 TAAACATCTCCACGCCCGCAGCAGGATTAATGTGCGCTTCATCTTCCAGGCGCAGGT 600  
 DB 541 TAAACATCTCCACGCCCGCAGCAGGATTAATGTGCGCTTCATCTTCCAGGCGCAGGT 600  
 QY 601 GAGCGTTGCTCCACCCCAAGAGCTACTCTGTTGAAACAGGAGGAAACCATGTGGATAG 660  
 DB 601 GAGCGTTGCTCCACCCCAAGAGCTACTCTGTTGAAACAGGAGGAAACCATGTGGATAG 660  
 QY 661 CGAATCTCGGTTAATGCGGATGTTAAACCGGATTTACGAAAGCGCTCCAGATAGTTGCGC 720  
 DB 661 CGAATCTCGGTTAATGCGGATGTTAAACCGGATTTACGAAAGCGCTCCAGATAGTTGCGC 720  
 QY 721 TTTAGTTTCTGCTGACGACACCATTTTCCGCGCTGCTGCAACAGGACTTCAACCGC 780  
 DB 721 TTTAGTTTCTGCTGACGACACCATTTTCCGCGCTGCTGCAACAGGACTTCAACCGC 780  
 QY 781 TTGCGTTGCTTTGCGGTTGGGTCGATACCAACACTCGAACCCAGCTGATGCTCGAG 840  
 DB 781 TTGCGTTGCTTTGCGGTTGGGTCGATACCAACACTCGAACCCAGCTGATGCTCGAG 840  
 QY 841 AGCTTTAACCGGCTGACTCAACCGCGAGGGGAAATGAAAAGGCTAAGGAGGCGGCTTC 900  
 DB 841 AGCTTTAACCGGCTGACTCAACCGCGAGGGGAAATGAAAAGGCTAAGGAGGCGGCTTC 900  
 QY 901 GAGCTGCTTCTCATCAATGATGAGGCAAGTCTCCAGTTCAATGGGGTTCAATGAAGCT 960  
 DB 901 GAGCTGCTTCTCATCAATGATGAGGCAAGTCTCCAGTTCAATGGGGTTCAATGAAGCT 960  
 QY 961 ATATTAAACCATGTTAAGAACCAATCATTTTACTTAAGTACTTCCATAGGTCCAGATGGT 1020  
 DB 961 ATATTAAACCATGTTAAGAACCAATCATTTTACTTAAGTACTTCCATAGGTCCAGATGGT 1020  
 QY 1021 GATCATGGAATCTTCATTAACAGTCTGCTTTTGGGGGCGAGTCTTTTACTGTTCATCGG 1080  
 DB 1021 GATCATGGAATCTTCATTAACAGTCTGCTTTTGGGGGCGAGTCTTTTACTGTTCATCGG 1080  
 QY 1081 ACCCGAGAGTACTGGTGATTAACAGGATTAAGCGGAGGACTCATTCCGGTCT 1140  
 DB 1081 ACCCGAGAGTACTGGTGATTAACAGGATTAAGCGGAGGACTCATTCCGGTCT 1140  
 QY 1141 TCTCGTGTGTTAAATTTCTGAGCTGTTTTGTTCATCGCGGACCTTGGGGGCTTGATCT 1200



Db	1141	TCTCGTGTGTTTAAATTTCTGAGCTCTTTTGTTCATCGCGCGGACCTTGGGGGTGGATCT	1200
QY	1201	TTTGTCCAAATGCGCGCGCATGCTGCTCGATATTAATTCGCTGGGGTGGCATGCGTTA	1260
Db	1201	TTTGTCCAAATGCCGGCGGATCGTGTCTGATTAATATGCGCTGGGGTGGCATGCTTA	1260
QY	1261	GTTATGGTTTGCCTCATATGACGAGGAAGAAGCCATGACAAACAAGGTGGAGCGCAC	1320
Db	1261	GTTATGGTTTGCCTCATATGACGAGGAAGAAGCCATGACAAACAAGGTGGAGCGCAC	1320
QY	1321	GATCATTAAGAAACAGAACCAACCGTSCCGCATGACACGCTTTGGCGGTTTCGGCG	1380
Db	1321	GATCATTAAGAAACAGAACCAACCGTSCCGCATGACACGCTTTGGCGGTTTCGGCG	1380
QY	1381	GGCCACTGACACGCGCAACCGGTTGCGGTGAGGTGAGCGTTCGTAAGCAGCGGGTTG	1440
Db	1381	GGCCACTGACACGCGCAACCGGTTGCGGTGAGGTGAGCGTTCGTAAGCAGCGGGTTG	1440
QY	1441	GGTAAAGCCCATGTTGATGGGAATCGTGTCTGACTGTTGAACCCCGAATGCGTATTG	1500
Db	1441	GGTAAAGCCCATGTTGATGGGAATCGTGTCTGACTGTTGAACCCCGAATGCGTATTG	1500
QY	1501	CGCGTTTGTGTTTATCGCGCGCTCGGCGCGCAATACGCGCACACCGCAGCGTGGAT	1560
Db	1501	CGCGTTTGTGTTTATCGCGCGCTCGGCGCGCAATACGCGCACACCGCAGCGTGGAT	1560
QY	1561	CGCGCTTGGCGGTTTCGCGGCAAGCCTGATCTGTGTTCCGCTGGTGGTTCGCGCCAG	1620
Db	1561	CGCGCTTGGCGGTTTCGCGGCAAGCCTGATCTGTGTTCCGCTGGTGGTTCGCGCCAG	1620
QY	1621	AGCATTTGACAGCCCGCTGTCACGCGCCAAAGTGTGGCGCTGGAATCAAGTCGTCG	1680
Db	1621	AGCATTTGACAGCCCGCTGTCACGCGCCAAAGTGTGGCGCTGGAATCAAGTCGTCG	1680
QY	1681	AGTTGTGATGACGCAATGGCCATCAAACTGATGTGAGTGGTTAGTTTCGCGGTTT	1740
Db	1681	AGTTGTGATGACGCAATGGCCATCAAACTGATGTGAGTGGTTAGTTTCGCGGTTT	1740
QY	1741	GAATTCGGTGGCTTCGCGCAAAATGTTGATGCGCGGCTCGTGGGAAATCATCGAT	1800
Db	1741	GAATTCGGTGGCTTCGCGCAAAATGTTGATGCGCGGCTCGTGGGAAATCATCGAT	1800
QY	1801	CTTCAACTCGCGCTCAGAAAATCTCAAGTTGTTGATGGAATCAAGGCTGTTGTCCAG	1860
Db	1801	CTTCAACTCGCGCTCAGAAAATCTCAAGTTGTTGATGGAATCAAGGCTGTTGTCCAG	1860
QY	1861	CTCAACTGACGAGACCAATCAATGACATGGTTCACGCTATCCGCGCGCTACTCT	1920
Db	1861	CTCAACTGACGAGACCAATCAATGACATGGTTCACGCTATCCGCGCGCTACTCT	1920
QY	1921	CTCGCGCAGCACCCATGCAAGCGCATCTCGCGAAGTGAATCGCCCGGTTCTGGG	1980
Db	1921	CTCGCGCAGCACCCATGCAAGCGCATCTCGCGAAGTGAATCGCCCGGTTCTGGG	1980
QY	1981	GTCATTGAGCTTGGGACCATATCAATATTTGTTTCAAGTTTCAACATGCCCTCAG	2040
Db	1981	GTCATTGAGCTTGGGACCATATCAATATTTGTTTCAAGTTTCAACATGCCCTCAG	2040
QY	2041	CTTACCTCGCTCGCGGGGAACCTCTGGAATTCATCGAGATATTGTTCGTCAG	2100
Db	2041	CTTACCTCGCTCGCGGGGAACCTCTGGAATTCATCGAGATATTGTTCGTCAG	2100
QY	2101	GCCTCGCAAGTGGTGAGAAAGCAATGAGCGAAGACCATGTTGTGGAGCTGACT	2160
Db	2101	GCCTCGCAAGTGGTGAGAAAGCAATGAGCGAAGACCATGTTGTGGAGCTGACT	2160
QY	2161	CAAGTTCTCAGCGTCATCGCGGTTCTCCACCAACGATTATGATGAATAGCTT	2220
Db	2161	CAAGTTCTCAGCGTCATCGCGGTTCTCCACCAACGATTATGATGAATAGCTT	2220
QY	2221	CTGATGAATCAGAGCGGCGGCTCTCTCGCCATGAATCAGCGCGCTCGCTGT	2280
Db	2221	CTGATGAATCAGAGCGGCGGCTCTCTCGCCATGAATCAGCGCGCTCGCTGT	2280

2281 CTCTGGACCGTAGGAGAAATCCACACGTAAAGAGCCCTTCCAGACGCAACAATGTCACG 2340  
 Db 2281 CTCTGGACCGTAGGAGAAATCCACACGTAAAGAGCCCTTCCAGACGCAACAATGTCACG 2340  
 2341 CAATGCGTACATGCTTTCTTCCAAAGGAGTATCT 2374  
 Db 2341 CAATGCGTACATGCTTTCTTCCAAAGGAGTATCT 2374  
 RESULT 2  
 A93933 2374 bp DNA linear PAT 22-JAN-2000  
 LOCUS Sequence 2 from Patent WO9723597.  
 DEFINITION A93933  
 ACCESSION A93933  
 VERSION A93933.1 GI:6742037  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Corynebacterium glutamicum  
 Corynebacterium glutamicum  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 1. (bases 1 to 2374)  
 REFERENCE Vrljić M. and Eggeling L.  
 AUTHORS  
 TITLE PROCESS FOR THE MICROBIAL PRODUCTION OF AMINO ACIDS BY BOOSTED  
 ACTIVITY OF EXPORT CARRIERS  
 JOURNAL Patent: WO 9723597-A 2 03-JUL-1997;  
 KERNFORSCHUNGSANLAGE JUELICH (DE); VRLJIĆ MARINA (DE)  
 FEATURES  
 source Location/Qualifiers  
 1. 2374  
 /organism="Corynebacterium glutamicum"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1718"  
 ORIGIN  
 Query Match 100.0%; Score 2374; DB 6; Length 2374;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 Qy 1 CCATTTCGTGAAGTGTTACTCTGCTGGCCCAATTCCTCGGCGGCAAGTGTGAAAAAC 180  
 Db 1 CCATTTCGTGAAGTGTTACTCTGCTGGCCCAATTCCTCGGCGGCAAGTGTGAAAAAC 180  
 Qy 61 CCTGAACCTTTTCAGAAGTAACCTAAGCCGCAATCCCTCGATTCTGCAATCAACGACGC 120  
 Db 61 CCTGAACCTTTTCAGAAGTAACCTAAGCCGCAATCCCTCGATTCTGCAATCAACGACGC 120  
 Qy 121 GTCTGTAGTCTACTAGATGATCTAGATTCACGGCGCCATCGTTGCGCAATACATCGGTGT 180  
 Db 121 GTCTGTAGTCTACTAGATGATCTAGATTCACGGCGCCATCGTTGCGCAATACATCGGTGT 180  
 Qy 181 GTCAATGGGTATCTCATCGAGGAGATCACTTCTCTGCTTTTAGCATGGGACGAGCTTG 240  
 Db 181 GTCAATGGGTATCTCATCGAGGAGATCACTTCTCTGCTTTTAGCATGGGACGAGCTTG 240  
 Qy 241 GGTTCGGGAAGAAGTCCCAACCAAGGCTTCGGCGAATTGCCTCACCAAAACCTTCGC 300  
 Db 241 GGTTCGGGAAGAAGTCCCAACCAAGGCTTCGGCGAATTGCCTCACCAAAACCTTCGC 300  
 Qy 301 CGACGGGCAATGATAGCGGCTCGCGCCCAAGGACCATCGACGCGGCCGTCCAGGTC 360  
 Db 301 CGACGGGCAATGATAGCGGCTCGCGCCCAAGGACCATCGACGCGGCCGTCCAGGTC 360  
 Qy 361 ACGGCTCTTGAAGCACAATCTTTGGGACCGAAGCGTAAAGACGGGCAATCGAGCTCATCTAG 420  
 Db 361 ACGGCTCTTGAAGCACAATCTTTGGGACCGAAGCGTAAAGACGGGCAATCGAGCTCATCTAG 420  
 Qy 421 TTTCGCCATCAACCAATGATGAGCATTCGCGCAATGAGGGGGTTGCAATGCCAAGTGCAGCAT 480  
 Db 421 TTTCGCCATCAACCAATGATGAGCATTCGCGCAATGAGGGGGTTGCAATGCCAAGTGCAGCAT 480  
 Qy 481 GGTTCGAAGTCTTACTACTTCAATCCGCGCAACGGGANTTAGCTTCAAGGTTACCGGCTCC 540  
 Db 481 GGTTCGAAGTCTTACTACTTCAATCCGCGCAACGGGANTTAGCTTCAAGGTTACCGGCTCC 540

QY	541	TAAGAATCTCCACGCGCGAGAGGATAATGTGTGCGCTTCATCTCCAGCGCAGCGT	600
Db	541	TAAGAATCTCCACGCGCGAGAGGATAATGTGTGCGCTTCATCTCCAGCGCAGCGT	600
QY	601	GAGCGTTGCTCCACCCCAAGAGCTACCTCGTTGAACACGCGGAGAAACATGTGATAG	660
Db	601	GAGCGTTGCTCCACCCCAAGAGCTACCTCGTTGAACACGCGGAGAAACATGTGATAG	660
QY	661	CGAATCTGCGCTGATGCGATGTAAAGGATTTGAGAGCGGCTCCAGATGTTGCG	720
Db	661	CGAATCTGCGCTGATGCGATGTAAAGGATTTGAGAGCGGCTCCAGATGTTGCG	720
QY	721	TTTAGTTCTGCTTCAGACACACATTTTCGGCGCTTGTGACAGAGACTTTCACCGC	780
Db	721	TTTAGTTCTGCTTCAGACACACATTTTCGGCGCTTGTGACAGAGACTTTCACCGC	780
QY	781	TTTCGGTTGCTTTGGCGGTTGGGTGCGGATACCAACACTCGACCCACGTTGCTCGAG	840
Db	781	TTTCGGTTGCTTTGGCGGTTGGGTGCGGATACCAACACTCGACCCACGTTGCTCGAG	840
QY	841	AGCTTTAACGCGTGACTCACCGCGAGGGGAAATGGAAAGGGCTAAGGAGCGCTTC	900
Db	841	AGCTTTAACGCGTGACTCACCGCGAGGGGAAATGGAAAGGGCTAAGGAGCGCTTC	900
QY	901	GAGCGTCTTCATCAATGATTGAGAGCAAGTGTCCAGTTGAATGGGTTTCATGAAGCT	960
Db	901	GAGCGTCTTCATCAATGATTGAGAGCAAGTGTCCAGTTGAATGGGTTTCATGAAGCT	960
QY	961	ATATTAACCACTGTTAAGCAACCAATCAATTTTACTTAAGTACTTTCATAGTCAAGTGT	1020
Db	961	ATATTAACCACTGTTAAGCAACCAATCAATTTTACTTAAGTACTTTCATAGTCAAGTGT	1020
QY	1021	GATCATGGAATCTTATCAAGCTCTGCTTTTGGGGCGAGTCTTTTATCTGCTCATCGG	1080
Db	1021	GATCATGGAATCTTATCAAGCTCTGCTTTTGGGGCGAGTCTTTTATCTGCTCATCGG	1080
QY	1081	ACCGCAGAAATGTAAGTAAACAGGAATTAAGCGCGAAGCACTCAATTCGGTCT	1140
Db	1081	ACCGCAGAAATGTAAGTAAACAGGAATTAAGCGCGAAGCACTCAATTCGGTCT	1140
QY	1141	TCCTGTTGTTTAAATTTGACGCTCTTTTGTTCATCGCGGCACTTGGGCTTGAATCT	1200
Db	1141	TCCTGTTGTTTAAATTTGACGCTCTTTTGTTCATCGCGGCACTTGGGCTTGAATCT	1200
QY	1201	TTTGTCTAATCGCGCGCGTCTGCTCGATATTATCGCTGGGTGGGATGCTTACCT	1260
Db	1201	TTTGTCTAATCGCGCGCGTCTGCTCGATATTATCGCTGGGTGGGATGCTTACCT	1260
QY	1261	GTTATGTTTGGCGTCATGGCAGAAAGACCGCATACACAAACAGGTGGAAGCGCCACA	1320
Db	1261	GTTATGTTTGGCGTCATGGCAGAAAGACCGCATACACAAACAGGTGGAAGCGCCACA	1320
QY	1321	GATCATGGAACAGAACCAACCGTGGCGGATGACAGCGCTTTGGCGGTTTCGGCGGT	1380
Db	1321	GATCATGGAACAGAACCAACCGTGGCGGATGACAGCGCTTTGGCGGTTTCGGCGGT	1380
QY	1381	GGCGCTGACACCGCGACCGGTTGGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG	1440
Db	1381	GGCGCTGACACCGCGACCGGTTGGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG	1440
QY	1441	GTTAAGCCCATGTTGATGCAATCGTGTGCTGACTGTTGAAACCCGATCGTATTGGA	1500
Db	1441	GTTAAGCCCATGTTGATGCAATCGTGTGCTGACTGTTGAAACCCGATCGTATTGGA	1500
QY	1501	CGCGTTTGTGTTTATCGCGCGTTCGCGCGCAATACGCGACCGCGGTGAGTATTT	1560
Db	1501	CGCGTTTGTGTTTATCGCGCGTTCGCGCGCAATACGCGACCGCGGTGAGTATTT	1560
QY	1561	CGCGCTGCGCGTTTCGCGCAAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620
Db	1561	CGCGCTGCGCGTTTCGCGCAAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620
QY	1621	ASCATTTGTCAGCGCGCTGCTCCAGCCCAAGTGTGAGTGAATCAACACGCTGCTGCTGCG	1680
Db	1621	ASCATTTGTCAGCGCGCTGCTCCAGCCCAAGTGTGAGTGAATCAACACGCTGCTGCTGCG	1680
QY	1681	AGTTGTGATGACCGATTTGGCCATCAAACTGATGTATGAGTGTAGTTTTCGCGGGTTT	1740
Db	1681	AGTTGTGATGACCGATTTGGCCATCAAACTGATGTATGAGTGTAGTTTTCGCGGGTTT	1740
QY	1741	GGAATCGTGTGCTTTCGCGAAATTTGATGCGCGCTGCTGGGAAATCTCATCGATCGC	1800
Db	1741	GGAATCGTGTGCTTTCGCGAAATTTGATGCGCGCTGCTGGGAAATCTCATCGATCGC	1800
QY	1801	CTCCACTGCGCTGACAGAAATCTCCAAAGTTTGTGAGTGAATCAAGGCTGTTGTCAGCTG	1860
Db	1801	CTCCACTGCGCTGACAGAAATCTCCAAAGTTTGTGAGTGAATCAAGGCTGTTGTCAGCTG	1860
QY	1861	CTCACTGACGAGACCAATCAATGACACTGCTCAAGTGTGAGTGAATCAAGGCTGTTGTCAGCTG	1920
Db	1861	CTCACTGACGAGACCAATCAATGACACTGCTCAAGTGTGAGTGAATCAAGGCTGTTGTCAGCTG	1920
QY	1921	CTCGCGACGACCCATGCAAGCGCATCTGCGCAAGTGAATGCTGCGCGCTTCTGCGCGAT	1980
Db	1921	CTCGCGACGACCCATGCAAGCGCATCTGCGCAAGTGAATGCTGCGCGCTTCTGCGCGAT	1980
QY	1981	GTCAATGAGCTTGGCGACCAATCAATATTTTTCAGTTCACATGCTCAACATGCCCTCAGACAGGA	2040
Db	1981	GTCAATGAGCTTGGCGACCAATCAATATTTTTCAGTTCACATGCTCAACATGCCCTCAGACAGGA	2040
QY	2041	CTTACCTGCTTGGCGGGAACCTCTGGAATTCATTCAGATATTTTTCGCTGAGGAG	2100
Db	2041	CTTACCTGCTTGGCGGGAACCTCTGGAATTCATTCAGATATTTTTCGCTGAGGAG	2100
QY	2101	GCCTTCGCGCAAGTGTGAGAAACCAATGACGCAAGCAATTTGTTGGCAGCTGACTGCAA	2160
Db	2101	GCCTTCGCGCAAGTGTGAGAAACCAATGACGCAAGCAATTTGTTGGCAGCTGACTGCAA	2160
QY	2161	CAAGTTCTCAGCGCTACCGCGCTTCTCCACCAAGTGAATTAATGATGAATGCTTGG	2220
Db	2161	CAAGTTCTCAGCGCTACCGCGCTTCTCCACCAAGTGAATTAATGATGAATGCTTGG	2220
QY	2221	CTGATGAATCAGAGCGGCGAGCTTCTCCGCAATGAACTCAGCGCGCTTCCGCTGAG	2280
Db	2221	CTGATGAATCAGAGCGGCGAGCTTCTCCGCAATGAACTCAGCGCGCTTCCGCTGAG	2280
QY	2281	CTCTGACCGTAGAGAAATACCCAGTAAGAGCTTTCCAGAGCAACAAATGTCAGC	2340
Db	2281	CTCTGACCGTAGAGAAATACCCAGTAAGAGCTTTCCAGAGCAACAAATGTCAGC	2340
QY	2341	CAATGCTGATGTTTCTTCCAAAGGATTAATCT	2374
Db	2341	CAATGCTGATGTTTCTTCCAAAGGATTAATCT	2374
RESULT 3			
AP05277/c			
LOCUS 333150 bp DNA linear BCT 08-AUG-2002			
DEFINITION Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 4/10			
ACCESSION AP05277 BA000036			
VERSION AP05277.1 GI:21323710			
KEYWORDS			
SOURCE			
ORGANISM			
Corynebacterium glutamicum ATCC 13032			
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
Corynebacteriaceae; Corynebacterium.			
REFERENCE			
1 Nakagawa, S.			
Complete genomic sequence of Corynebacterium glutamicum ATCC 13032			
Unpublished			
2 (bases 1 to 333150)			
REFERENCE			
Nakagawa, S.			
Direct Submission			
Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.			

Thu Mar 18 12:30:58 2004

Ind., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,  
Tokyo 194-8533, Japan (E-mail: snakagawa@anagen.com,  
Tel: 81-44-829-3031, Fax: 81-44-813-1651)  
This sequence is conducted by collaboration of Kyowa Hakko Kogyo  
Co. Ltd. and Kitagato University.

## FEATURES

Location/Qualifiers  
1..333150  
/organism="Corynebacterium glutamicum ATCC 13032"  
/mol\_type="genomic DNA"  
/strain="ATCC 13032"  
/db\_xref="taxon:196627"  
/note="ATCC 13032"  
complement(63..1292)  
/gene="Cg10944"  
complement(63..1292)  
/gene="Cg10944"  
/codon\_start=1  
/transl\_table=11  
/product="Hypothetical membrane protein"  
/protein\_id="BAB98337.1"  
/db\_xref="GI:21323711"  
/translation="MTNPTTEARNRRLIWNLQNGDQIVAAKTVLPMLQAQAPG  
FILALLPIREAGSMPLQAAITGMVLRQTSKVMVIGSQFVSALGIGVRLAFLRG  
WALGITVLLAALSLFRSMCSIAKDVQKGVSKGRGLVTRATVIGGWLVAFLG  
ATAFLGSHSPVLAANVTASSPSMIFASIVFARIEYAKPATPKNAPSANPWVRCCI  
AAUKDKAFRRFVLVRSMLVTALSTAFILVALESNSIDSLGFLIASGLASLVGGG  
RISGWSHSSKNWAGGALFGSIVLILVLSNAPAPAINTLVPLISFLILTAHTA  
IRVARKTVMMAGDQRTYRVADANTLMVLLIVGALSGFIALFGNEAALLFLAAI  
GLGITSARGLKEVSAG"  
1382..2917  
/gene="Cg10945"  
1382..2917  
/gene="Cg10945"  
/codon\_start=1  
/transl\_table=11  
/product="Putative multicopper oxidases"  
/protein\_id="BAB98338.1"  
/db\_xref="GI:21323712"  
/translation="MPELNRRFFKAGVLAATVGAQVLVACSSDDVGVGGEPRTL  
PIPPADLTREGSVFHALPEAQGESQILPDVTKTWGFNGHLGPTLVUKGDDVHV  
DVNLNDEMTVHMGKMLPAIADGPHSPIGQGTWSTVTVANDAAITLHAPHTHG  
LTGLHVRGLAGMIIVDEADTKLDLPREYGVDDIPLVMDHRFLEDGSLDEBDDPL  
GLIGDPTANGITNAHPDATTRVRFRVLNGSNMRFLAFSDTRTFQVLIASGSLD  
EPQRTTLAGGERWEIVELPEGDVTLSSGVFDNYGVDPDEFVDFQMSDSFOL  
LTITGPSDAAQAPALPGVLKSTEDVDIDATERFTIMTFESINDLQMDORVDVVID  
HDQPEVITVNDNSDPNHFHDAFVKLPFGEDVDELFGNCKDTVGLPPGATATL  
AVEFGHYPPQWPMYMYHCHMLYHEDQGMWGQFVIVPEGDEPAALVGSSTGSSIDSAGG  
HAH"  
3063..3983  
/gene="Cg10946"  
3063..3983  
/gene="Cg10946"  
/note="PF00005:ABC transporter  
TIGR00960:3a0501s02: Type II (General) Secretory Pathway  
(TISP) Family protein  
TIGR00968:3a0106s01: sulfate transport system permease  
protein  
TIGR01166:cbiO: cobalt transport protein ATP-binding  
subunit  
TIGR01184:ntrCD: nitrate transport ATP-binding subunits C  
and D  
TIGR01186:proV: glycine betaine/L-proline transport ATP  
binding subunit  
TIGR01187:potA: spermidine/putrescine ABC transporter  
ATP-binding subunit  
TIGR01188:draA: daunorubicin resistance ABC transporter  
ATP-binding subunit  
TIGR01189:ccmA: heme exporter protein CcmA"  
/codon\_start=1  
/transl\_table=11  
/product="ABC-type transporter, ATPase component"  
/protein\_id="BAB98339.1"  
/db\_xref="GI:21323713"

/translation="MTTTPAIDVTLYRTYGYTAVKGLNPHVORGVEFGLLGTNGAG  
KXTSLVEIGLSAPSGTVRISGDPVADRAILRDELGLMLOSGLPSQLTVAETMDM  
WGTCTVPRAIKDVLADVLDHRENVKVGALSGEQRIDILACALIGDPSILFLEPT  
TGLDSESRITWQLLDLKQKQVMTMTTHYLEEAEFLCDRIAIMNAGEIAGEGILDE  
LVAREKSTISFLRGQVELPVLGABIIIRDNNHVRITATTTTQQHTLEILTAAETGI  
ALEGFAKPAATLESVFMDSIASLENTSLQTA"  
4003..4836  
/gene="Cg10947"  
4003..4836  
/gene="Cg10947"  
/codon\_start=1  
/transl\_table=11  
/product="ABC-type transporter, permease components"  
/protein\_id="BAB98340.1"  
/db\_xref="GI:21323714"  
/translation="MTTSHARGLOHATPERRKTSFFKTSFKABWLOFRNKTLIFM  
ATVFPVGIPLLFLINGGAASANSFDFVMYTLIFVQFVTSMAITRRDRSVLR  
LRTGARIDIIGAIKCPGALLTLFTVVIPLIMVLGAPINIPVIFAVILGILL  
CSALUMSGATRNAEAQMTSMVPMGLMGSLGSRFVFGDSIVADILATPFAAIS  
DLVQIGWAGTADSVGSGVERANFAGIFQDMLIPGLILAATAAANAANRYMRDYSY  
R"  
4879..6138  
/gene="Cg10948"  
4879..6138  
/gene="Cg10948"  
/EC\_number="2.7.3.-"  
/codon\_start=1  
/transl\_table=11  
/product="Two-component system, sensory transduction  
histidine kinases"  
/protein\_id="BAB98341.1"  
/db\_xref="GI:21323715"  
/translation="WNKDFWAGTARWFSRGVSLASPVTAFLNSRRRLNLAKYTL  
YTRVSLQAIIPVVLISLAFGLIVANAGTLNPFVWLLGFSVILLIVTVLVYQPSLNS  
HPRSVQPPFFTGLVNLGVVSVVQLPGLMNSDNRATLIFTICVFLLSIAYI  
PMNYRWMLIAMSALVMTSTTDLGSLAWVIPPMLMAGTVLSVTVTVVMKVEKRS  
RELEASLVTERLFAQLHDLTGLQHLAAMSVELAKRGGDRLENELELOK  
LRTISMSEMDVSGYRTVNLATEISGAKSLIADAHLSVIGTTSQVSPAHRELCAW  
LVREATTNLRSLDAITLTLSTSTVMDNNGVKNKIGRLSGLSALRSASAGMTL  
IVSRDDQFSVRMLINAPANTPAEKA"  
6139..6750  
/gene="Cg10949"  
6139..6750  
/gene="Cg10949"  
/note="PF00072:Response regulator receiver domain  
PF00196:Bacterial regulatory proteins, luxR family"  
/codon\_start=1  
/transl\_table=11  
/product="Two-component system, response regulators  
consisting of a CheY-like receiver domain and a HTH  
DNA-binding domain"  
/protein\_id="BAB98342.1"  
/db\_xref="GI:21323716"  
/translation="MISISADDEALIASLATLSLEPDLDPVTPAGSGEELIETWA  
DPSNTDVCVLQGLQIGDIDTATRLMETTPDLAVITVSHARPKOLKALAAVGLG  
FLPKTSADFEATARTVHAGRYIDPELAAMTISAGESPLTNRREEVLELAQGLSA  
ESIAVAHAHAPGTRNYLSQAMTKVGAQNFEEATFRELQGL"  
complement(6781..6987)  
/gene="Cg10950"  
complement(6781..6987)  
/gene="Cg10950"  
/codon\_start=1  
/transl\_table=11  
/product="Hypothetical protein"  
/protein\_id="BAB98343.1"  
/db\_xref="GI:21323717"  
/translation="MKHYMSYFGAGPLNTPTHTKEYVFMFDIVSVINHFWEKSTNNWVN  
WGLPREFVPAFVGSTVLVLSF"  
7048..8205  
/gene="Cg10951"  
7048..8205  
/gene="Cg10951"  
/codon\_start=1



FEATURES		Location/Qualifiers	
source		1. 349980	
		/organism="Corynebacterium glutamicum"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:1718"	
		/note="Seq 1 to long (3.309.400) split in 11, seq 7063	
		1.200.001 1.549.980"	
ORIGIN		Query Match	
		100.0%; Score 2374; DB 6; Length 349980;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 2374; Conservative 0; Mismatches 0; Gaps 0;	
		1 CCAATTTCTGAGAGTGTACTCTGCTGCGCCCAATTCCTGCGGCGCAAGAGTGAAGAAAC 60	
Qy		CCATTTGCTGAAAGTGTGTACTCTGCTTGGCCATTCCTGCGGCGCAAGAGTGAAGAAAC 129909	
Db		CCTGAACCTTTTCAGAAAGTAACTTAAGGCGCGCAATCCCTCGATTTGTCATCAACGAGCGC 120	
Qy		61 CCTGAACCTTTTCAGAAAGTAACTTAAGGCGCGCAATCCCTCGATTTGTCATCAACGAGCGC 129849	
Db		121 GTCTGTGAGTCTAGCTAGAGATCTAGATTCAGGCGCATCTGTCCTCAATACATCGGTGT 180	
Qy		129848 GTCTGTGAGTCTAGCTAGAGATCTAGATTCAGGCGCATCTGTCCTCAATACATCGGTGT 129789	
Db		181 GTCAATGGGTATCTCATCGAGAGATCACTTCTCTGCTTTTAGCATGGGAGAGCTTG 240	
Qy		129788 GTCAATGGGTATCTCATCGAGAGATCACTTCTCTGCTTTTAGCATGGGAGAGCTTG 129729	
Db		241 GGTTCGCGGGAAGAGTCCCAACCAAGGGCTCGGGCAATTCCTCACCAAAACCTTCGCG 300	
Qy		129728 GGTTCGCGGGAAGAGTCCCAACCAAGGGCTCGGGCAATTCCTCACCAAAACCTTCGCG 129669	
Db		301 CGAGCGGCAATGGATGATCGCGCTGCGCCCAAGAGCATTCAGACGCGCGCTCCAGGTC 360	
Qy		129668 CGAGCGGCAATGGATGATCGCGCTGCGCCCAAGAGCATTCAGACGCGCGCTCCAGGTC 129609	
Db		361 AGGCTCTGGAAGCACATCTTTGGGACCGAAGGTAAAGCGGCAATCGAGCGCATCTAG 420	
Qy		129608 AGGCTCTGGAAGCACATCTTTGGGACCGAAGGTAAAGCGGCAATCGAGCGCATCTAG 129549	
Db		421 TTTCCATCAACCATGTAGCGCATTCGCGCATGAGGGGTTCATGCGCAAGTGGCGCAT 480	
Qy		129548 TTTCCATCAACCATGTAGCGCATTCGCGCATGAGGGGTTCATGCGCAAGTGGCGCAT 129489	
Db		481 GGTTCGCAAGTTCTACTTCTCATCCGCGCAAGGATTTAGCTTCAGGGTTCACGCTCC 540	
Qy		129488 GGTTCGCAAGTTCTACTTCTCATCCGCGCAAGGATTTAGCTTCAGGGTTCACGCTCC 129429	
Db		541 TAAACATCTCCACGCGCAGCAAGGATAATGTGTGCGCTTCACTTCCAGCGCAGCGT 600	
Qy		129428 TAAACATCTCCACGCGCAGCAAGGATAATGTGTGCGCTTCACTTCCAGCGCAGCGT 129369	
Db		601 GAGCGTTGCTCCACCCCAAGAGTACTCTGTTGACACGCGGAGAAACCATGTGATAG 660	
Qy		129368 GAGCGTTGCTCCACCCCAAGAGTACTCTGTTGACACGCGGAGAAACCATGTGATAG 129309	
Db		661 CGAATCTGCGTTGATGCGGATGTTTAAAGCGATTTTCAAGACGCGCTCCAGATAGTGGC 720	
Qy		129308 CGAATCTGCGTTGATGCGGATGTTTAAAGCGATTTTCAAGACGCGCTCCAGATAGTGGC 129249	
Db		721 TTTAGTTTCTGCTTGCAGCAACACCAATTTTCCGCGTGTCTGCAAGAGCTTCACCCGC 780	
Qy		129248 TTTAGTTTCTGCTTGCAGCAACACCAATTTTCCGCGTGTCTGCAAGAGCTTCACCCGC 129189	
Db		781 TTCCGTTGCTTTGGCCGGTGGGTGGCGATACCACTCCGACCCAGCTGATCTCGAG 840	
Qy		129188 TTCCGTTGCTTTGGCCGGTGGGTGGCGATACCACTCCGACCCAGCTGATCTCGAG 129129	
Db		841 AGCTTTAAGCGCTGACTCACCGCGGAGGGGAAATGGAAGGGCTTAAGAGGGCGCTTC 900	
Qy		129128 AGCTTTAAGCGCTGACTCACCGCGGAGGGGAAATGGAAGGGCTTAAGAGGGCGCTTC 129069	
Db			

328458	CGCGCTGGCGCTTCGCGCAAGCCGTATGTGTTCCGCTGTGGTTCGCGCAGC	328399
Qy	1621 AGCATTTGTCAGCGCCCTCTCGAGCCCAAGAGTGTGGCGCTGGATCAACGTCGTGGC	1680
Db	328398 AGCATTTGTCAGCGCCCTCTCGAGCCCAAGAGTGTGGCGCTGGATCAACGTCGTGGC	328339
Qy	1681 AGTTGTGATGACCGCAATTCGCGCATCAACTGATGTGTGATGGTTAGTTTCGCGGTTT	1740
Db	328338 AGTTGTGATGACCGCAATTCGCGCATCAACTGATGTGTGATGGTTAGTTTCGCGGTTT	328279
Qy	1741 GGAATCGGTGGCTTCGCGCAATTCGCGCATCAACTGATGTGTGATGGTTAGTTTCGCGGTTT	1800
Db	328278 GGAATCGGTGGCTTCGCGCAATTCGCGCATCAACTGATGTGTGATGGTTAGTTTCGCGGTTT	328219
Qy	1801 CTCCAACTCGCGCTCAGAAACTTCCAAAGTGTGTGATGAAATCAAGCGTGTTCAGCTG	1860
Db	328218 CTCCAACTCGCGCTCAGAAACTTCCAAAGTGTGTGATGAAATCAAGCGTGTTCAGCTG	328159
Qy	1861 CTCCAACTGACGAAGCACCAATCAATGCACTGGTCACTGATCCGCGCTACTCTCTTG	1920
Db	328158 CTCCAACTGACGAAGCACCAATCAATGCACTGGTCACTGATCCGCGCTACTCTCTTG	328099
Qy	1921 CTGCGCAGCAGCCATCGAGCGCATCTGCGCAAGTGAAGTCCGCGCTTCTGCGCGAT	1980
Db	328098 CTGCGCAGCAGCCATCGAGCGCATCTGCGCAAGTGAAGTCCGCGCTTCTGCGCGAT	328039
Qy	1981 GTCAATGAGCTTTCGCGACCATATCAATATTGTTCACTTCAAGTGCCTCAGACAGGA	2040
Db	328038 GTCAATGAGCTTTCGCGACCATATCAATATTGTTCACTTCAAGTGCCTCAGACAGGA	327979
Qy	2041 CTTACCTCGTGGCGGGGAAACCTCTGGAATTCATGAGATATTTGTCGCTGAGCAG	2100
Db	327978 CTTACCTCGTGGCGGGGAAACCTCTGGAATTCATGAGATATTTGTCGCTGAGCAG	327919
Qy	2101 GCGTGGCAAGTGTGAGAAGCAATGACGCGCAAGACCATTTGTCGAGCTGACTGCA	2160
Db	327918 GCGTGGCAAGTGTGAGAAGCAATGACGCGCAAGACCATTTGTCGAGCTGACTGCA	327859
Qy	2161 CAAGTCTCTACCGCTCATCGCGCGTTCCTCCACCAAGATTAATGATGGAATAGCTGG	2220
Db	327858 CAAGTCTCTACCGCTCATCGCGCGTTCCTCCACCAAGATTAATGATGGAATAGCTGG	327799
Qy	2221 CTGATGATCAGAAGCGGCGAGCTCTCTCGCGATGACTCAGCGCGCTCGCTGTGAG	2280
Db	327798 CTGATGATCAGAAGCGGCGAGCTCTCTCGCGATGACTCAGCGCGCTCGCTGTGAG	327739
Qy	2281 CTCTGGACCGTAGGAAGAAATACCGAGTAAAGAGCCTTCCAGACGCAACATGTACG	2340
Db	327738 CTCTGGACCGTAGGAAGAAATACCGAGTAAAGAGCCTTCCAGACGCAACATGTACG	327679
Qy	2341 CAATCGGTACATGTTTCTTCCAAAGAGTATCT	2374
Db	327678 CAATCGGTACATGTTTCTTCCAAAGAGTATCT	327645

RESULT 4	
AX127147/c	349980 bp DNA linear PAT 11-MAY-2001
LOCUS	AX127147
DEFINITION	Sequence 7063 from Patent EP1108790.
ACCESSION	AX127147 AX1.4121
VERSION	AX127147.1 GI:14041135
KEYWORDS	Corynebacterium glutamicum
SOURCE	Corynebacterium glutamicum
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE	1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
AUTHORS	Novel polynucleotides
TITLE	Patent: EP 1108790-A 7063 20-JUN-2001;
JOURNAL	KYOWA HAKKO KOGYO CO., LTD. (JP)

Thu Mar 18 12:30:58 2004

QY	901	GAAGCTGCTTCATCAATGATTTGAGAGCAAGTGTCCAGTTGAAATGGGGTTCAAGAGCT	960	127988	GTCAATGAGCTTCCGACCATATCAATATTTGTTACGTTCAACATGCCCTCAGACAGGGA	127929	
Db	129068	GAAGCTGCTTCATCAATGATTTGAGAGCAAGTGTCCAGTTGAAATGGGGTTCAAGAGCT	129009	2041	CTTACCCCTGGCTGGCGGGAAACCCCTCTCGAATTTCCATCCGAGATATTTCTTCGGTACGACAG	2100	
QY	961	ATATTAAACCAATGTTAGAACCAATCATTTTACTTAAAGTACTTCCATAGTGCAGATGGT	1020	127928	CTTACCCCTGGCTGGCGGGAAACCCCTCTCGAATTTCCATCCGAGATATTTCTTCGGTACGACAG	127869	
Db	129008	ATATTAAACCAATGTTAGAACCAATCATTTTACTTAAAGTACTTCCATAGTGCAGATGGT	128949	2101	GCCCTGGCGAAGTGGTGAGAAAGCAATGACGCCCAAGACCAATGTTGGCAGCTGACTGCACA	2160	
QY	1021	GATCATGGAATCTTCATACAGTCTGCTTTTGGGGGGCCAGTCTTTTACTGTGCCATCGG	1080	127868	GCCCTGGCGAAGTGGTGAGAAAGCAATGACGCCCAAGACCAATGTTGGCAGCTGACTGCACA	127809	
Db	128948	GATCATGGAATCTTCATACAGTCTGCTTTTGGGGGGCCAGTCTTTTACTGTGCCATCGG	128889	2161	CAAGTTCTCACCGTCATCCGCCGTTCTCCACCAAGGATTAATGATGGAATAGCTTGG	2220	
QY	1081	ACCGCAGAACTGCTGGTGAATAAACAAGGAATTAAGCGGAGAGACTCATTTGGCGTTCT	1140	127808	CAAGTTCTCACCGTCATCCGCCGTTCTCCACCAAGGATTAATGATGGAATAGCTTGG	127749	
Db	128888	ACCGCAGAACTGCTGGTGAATAAACAAGGAATTAAGCGGAGAGACTCATTTGGCGTTCT	128829	2221	CTGATGAATCAGAAGCGGGGAGCGCTCTCTCGCATGAACCTCAGCGCTCCGCTGTGAG	2280	
QY	1141	TCTGCTGTGTTAAATTTCTGACGCTCTTTTGTTCATCGCGGCACTTGGCGGCTTGAATCT	1200	127748	CTGATGAATCAGAAGCGGGGAGCGCTCTCTCTCGCATGAACCTCAGCGCTCCGCTGTGAG	127689	
Db	128828	TCTGCTGTGTTAAATTTCTGACGCTCTTTTGTTCATCGCGGCACTTGGCGGCTTGAATCT	128769	2281	CTCTGGACCGTAGAAGAAATACCCAGTAAAGAGCTTTTCCAGACGCAACAATGTACAG	2340	
QY	1201	TTTGTCCAAATGCGCGCCGATCGTCTCGATATTAATGCGCTGGGGTGGCATCGCTTACCT	1260	127688	CTCTGGACCGTAGAAGAAATACCCAGTAAAGAGCTTTTCCAGACGCAACAATGTACAG	127629	
Db	128768	TTTGTCCAAATGCGCGCCGATCGTCTCGATATTAATGCGCTGGGGTGGCATCGCTTACCT	128709	2341	CAATGGGTACATGCTTCTTCCAAAGGAGTATCT	2374	
QY	1261	GTATGCTTTCCTGTCATGCGCAGAAAGACGATGACAAACAGGTGAAGCGGCACA	1320	127628	CAATGGGTACATGCTTCTTCCAAAGGAGTATCT	127595	
Db	128708	GTATGCTTTCCTGTCATGCGCAGAAAGACGATGACAAACAGGTGAAGCGGCACA	128649	RESULT 5			
QY	1321	GATCAATGAGAAACAGAACCAACGTCGCGATGACAGCTTTGGCGGTTTGGCGGT	1380	AX063767/c	AX063767	993 bp DNA linear PAT 24-JAN-2001	
Db	128648	GATCAATGAGAAACAGAACCAACGTCGCGATGACAGCTTTGGCGGTTTGGCGGT	128589	LOCUS	Sequence 49 from Patent WO0100843.		
QY	1381	GGCCACTGACAGCGCAACCGGTGGGGTGGAGTGCAGTGCATAGCAGCGGTTTG	1440	DEFINITION	AX063767		
Db	128588	GGCCACTGACAGCGCAACCGGTGGGGTGGAGTGCAGTGCATAGCAGCGGTTTG	128529	ACCESSION	AX063767.1	GI:12541479	
QY	1441	GGTAAAGCCCATGTTGATGGAATCTGCTGACCTGTTGAAACCGAATGCGATTGGA	1500	VERSION			
Db	128528	GGTAAAGCCCATGTTGATGGAATCTGCTGACCTGTTGAAACCGAATGCGATTGGA	128469	KEYWORDS		Corynebacterium glutamicum	
QY	1501	CGCGTTTGTGTTTATCGCGCGCTGCGCGCAATAGCGGACACCGGAGTGGATTTT	1560	SOURCE		Corynebacterium glutamicum	
Db	128468	CGCGTTTGTGTTTATCGCGCGCTGCGCGCAATAGCGGACACCGGAGTGGATTTT	128409	ORGANISM		Corynebacterium glutamicum	
QY	1561	CGCGCTGCGCGCTTCCGCGCAACGCTCATCTGTTCCCGCTGGTGGGTTTCGCGCAGC	1620	REFERENCE		1 Pompejus M., Kroeger, B., Schroeder, H., Zelder, O. and Haberhauer, G.	
Db	128408	CGCGCTGCGCGCTTCCGCGCAACGCTCATCTGTTCCCGCTGGTGGGTTTCGCGCAGC	128349	AUTHORS		corynebacterium glutamicum genes encoding metabolic pathway	
QY	1621	AGCAATGTCACGCGCGCTGTCACGCCCAAGTGTGGCTGGATCAAGTGTGCTGGC	1680	TITLE		Proteins	
Db	128348	AGCAATGTCACGCGCGCTGTCACGCCCAAGTGTGGCTGGATCAAGTGTGCTGGC	128289	JOURNAL		Patent: WO 0100843-A 49 04-JAN-2001;	
QY	1681	AGTTGTGATGACCGCATGCGCATCAACTGATGTTGATGGTTAGTTTTCGCGGTTT	1740	FEATUERS		Location/Qualifiers	
Db	128288	AGTTGTGATGACCGCATGCGCATCAACTGATGTTGATGGTTAGTTTTCGCGGTTT	128229	source		1. .993	
QY	1741	GGAACTGCTGCGCTTCGCCCAATGTTGATGCGCGGCTGCTGGAAATCTCATCGATCG	1800	CDS		/organism="Corynebacterium glutamicum"	
Db	128228	GGAACTGCTGCGCTTCGCCCAATGTTGATGCGCGGCTGCTGGAAATCTCATCGATCG	128169			/mol_type="unassigned DNA"	
QY	1801	CTCCAACTGCGGCTCAGAAACTCCCAAGTGTGTTGATGAACTCAAGGCTGTTGTCAGCTG	1860			/db_xref="taxon:1718"	
Db	128168	CTCCAACTGCGGCTCAGAAACTCCCAAGTGTGTTGATGAACTCAAGGCTGTTGTCAGCTG	128109			101. .973	
QY	1861	CTCAACTGACGAGCACAATCAATGCACTGTCAGCGTATCCGCGGCTACTCTCTTG	1920	ORIGIN		/notes="unnamed protein product; RXA01393"	
Db	128108	CTCAACTGACGAGCACAATCAATGCACTGTCAGCGTATCCGCGGCTACTCTCTTG	128049	Query Match		41.8%; Score 993; DB 6; Length 993;	
QY	1921	CTCGCAGCAGCCCATGCAAGCGCATCTGCGCAAGTACTGCGCGGCTCTTGGCGCAT	1980	Best Local Similarity		100.0%; Pred. No. 5.4e-233;	
Db	128048	CTCGCAGCAGCCCATGCAAGCGCATCTGCGCAAGTACTGCGCGGCTCTTGGCGCAT	127989	Matches 993; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1981	GTCAATGAGCTTGGGACCATATCAATATTTGTTCACTTCAACATGCTCCAGACAGGGA	2040	QY	62	CTGAACCTTTTCAGAGTAATTAAGCGCGCAATCCCTCGATTGCTGCATCAAGCGCGG	121
Db				Db	993	CTGAACCTTTTCAGAGTAATTAAGCGCGCAATCCCTCGATTGCTGCATCAAGCGCGG	934





QY 662 GAATCTGCTTGGATGCGATGCTTAAAGGGAATTTACAGGAGCGTCCAGATAGTTGCGCT 721  
Db 393 GAACTCTGCTTGAATGCGATGCTTAAAGGGAATTTACAGGAGCGTCCAGATAGTTGCGCT 334  
QY 722 TTAGTTTCTGCTTGAGGACACACATTTTCGCGCTGCTTGACAAAGGACTTCAACCGCT 781  
Db 333 TTAGTTTCTGCTTGAGGACACACATTTTCGCGCTGCTTGACAAAGGACTTCAACCGCT 274  
QY 782 TCGGTTCTTTCGCGCTTGGGTCGCGGATACCAACATCGACCCAGCTGATGCTCGAGA 841  
Db 273 TCGGTTCTTTCGCGCTTGGGTCGCGGATACCAACATCGACCCAGCTGATGCTCGAGA 214  
QY 842 GCTTTACGCGCTGACTCAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 901  
Db 213 GCTTTACGCGCTGACTCAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 154  
QY 902 AAGCTGCTTTCATCAATGATGAGAGCAAAAGTGTCCAGTTGAATGGGTTTCATGAAGCTA 961  
Db 153 AAGCTGCTTTCATCAATGATGAGAGCAAAAGTGTCCAGTTGAATGGGTTTCATGAAGCTA 94  
QY 962 TATTAACACATGTTAAGACCAATCATTTTACTTAACTACTTCCATAGGTCAGATGGTG 1021  
Db 93 TATTAACACATGTTAAGACCAATCATTTTACTTAACTACTTCCATAGGTCAGATGGTG 34  
QY 1022 ATCATGGAATCTTCATTAACAGGTCGCTTTTG 1054  
Db 33 ATCATGGAATCTTCATTAACAGGTCGCTTTTG 1  
RESULT 7  
AX123540/c  
LOCUS AX123540 870 bp DNA linear PAT 11-MAY-2001  
DEFINITION Sequence 3456 from Patent EPI108790.  
ACCESSION AX123540  
VERSION AX123540.1 GI:14041028  
KEYWORDS  
SOURCE Corynebacterium glutamicum  
ORGANISM Corynebacterium glutamicum  
Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
REFERENCE 1  
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.  
TITLE Novel polynucleotides  
JOURNAL Patent: EP 1108790-A 3456 20-JUN-2001;  
KYOWA HAKKO KOGYO CO., LTD. (JP)  
FEATURES  
source 1..870  
Location/Qualifiers  
/organism="Corynebacterium glutamicum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1718"  
ORIGIN  
Query Match 36.6%; Score 870; DB 6; Length 870;  
Best Local Similarity 100.0%; Pred. No. 9.5e-203;  
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 85 AGCGCGAATCCCTGATGCTGCTCAACAGGAGCGCTGCTGAGTCTAGCTAGATCT 144  
Db 870 AGCGCGAATCCCTGATGCTGCTCAACAGGAGCGCTGCTGAGTCTAGCTAGATCT 811  
QY 145 AGATTCAGGCGCATGCTGCTGCTCAACATCATCGGTGTGCTCAATGGGTATCTCATCGAGGAG 204  
Db 810 AGATTCAGGCGCATGCTGCTGCTCAACATCATCGGTGTGCTCAATGGGTATCTCATCGAGGAG 751  
QY 205 GATCACTTCTCTCTTTTACGATGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 264  
Db 750 GATCACTTCTCTCTTTTACGATGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691  
QY 265 AAGGCTCGGGAATGCTTCAACAAACCTTCGCCGAGGAGCAATGGATAGCGGCT 324  
Db 690 AAGGCTCGGGAATGCTTCAACAAACCTTCGCCGAGGAGCAATGGATAGCGGCT 631

QY 325 GCGCCCCACAGGACCATCGAGCGCGCCGCTCCAGGTCACGCTCTTGAGACATCTTTGGG 384  
Db 630 GCGCCCCACAGGACCATCGAGCGCGCCGCTCCAGGTCACGCTCTTGAGACATCTTTGGG 571  
QY 385 ACCGAAGGTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 444  
Db 570 ACCGAAGGTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 511  
QY 445 CCGCAATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 504  
Db 510 CCGCAATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 451  
QY 505 TCGCGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 564  
Db 450 TCGCGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 391  
QY 565 GGATAATGTTGCTGCTTTCATCTTCCAAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 624  
Db 390 GGATAATGTTGCTGCTTTCATCTTCCAAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 331  
QY 625 TACCTCGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 684  
Db 330 TACCTCGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 271  
QY 685 TAAAGG 744  
Db 270 TAAAGG 211  
QY 745 CATTTTCGCGCTGCTTTCATCTTCCAAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 804  
Db 210 CATTTTCGCGCTGCTTTCATCTTCCAAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 151  
QY 805 GCGGAGTACACACATCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 864  
Db 150 GCGGAGTACACACATCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 91  
QY 865 CGAGG 924  
Db 90 CGAGG 31  
QY 925 GAGCAAGAGTCTCAGTTGAATGGGTTTCAAT 954  
Db 30 GAGCAAGAGTCTCAGTTGAATGGGTTTCAAT 1  
RESULT 8  
BD165657/c  
LOCUS BD165657 870 bp DNA linear PAT 17-JAN-2003  
DEFINITION Novel polynucleotide.  
ACCESSION BD165657  
VERSION BD165657.1 GI:27871469  
KEYWORDS JP 2002191370-A/3456.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 870)  
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.  
TITLE Novel polynucleotide  
JOURNAL Patent: JP 2002191370-A 3456 09-JUL-2002;  
COMMENT OS Corynebacterium glutamicum  
PN JP 2002191370-A/3456  
PD 09-JUL-2002  
PF 15-DEC-2000 JP 2000405096  
PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,  
PI KEIKO OCHIAI,  
PI HARUHIKO YOKOI, NAOKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO  
PI OZAKI  
PC C12N15/09, C12N15/09, C07K16/12, C07K16/12, C07K16/40, C12M1/00, PC  
C12N1/15,  
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/



PC 04.C12P13/08,  
PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53, PC  
G01N33/566,  
PC G01N33/569,G01N33/68,G01N37/00//C12P21/08,(C12N1/21,C12R1:15),  
PC (C12N1/21,C12R1:13),(C12N1/21,C12R1:01),(C12P13/08,C12R1:15),  
PC C12N15/00,  
PC C12N5/00,C12N15/00  
CC Novel polynucleotide  
PH Key Location/Qualifiers  
FT source 1..870  
FT /organism='Corynebacterium glutamicum'.  
FEATURES  
source  
Location/Qualifiers  
1..870  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 36.6%; Score 870; DB 6; Length 870;  
Best Local Similarity 100.0%; Pred.No.9.5e-203; Indels 0; Gaps 0;  
Matches 870; Conservative 0; Mismatches 0;  
QY 85 AGCGCGAATCCCTCGATTGCTGCAATCAAGAGCGGCTCTGTGAGTCTAGTAGATCT 144  
Db |  
QY 870 AGCGCGAATCCCTCGATTGCTGCAATCAAGAGCGGCTCTGTGAGTCTAGTAGATCT 811  
Db |  
QY 145 AGATTCCAGCGGCATCGTTGCCAATACATCGGTGTCAATGGGTATCTCATCGAGAG 204  
Db |  
QY 810 AGATTCCAGCGGCATCGTTGCCAATACATCGGTGTCAATGGGTATCTCATCGAGAG 751  
Db |  
QY 205 GATCACTTCTCTCTCTTTAGTAGTGGAGCAGCTTCGGTTTCGGGAGAAAGTCCCCAAC 264  
Db |  
QY 750 GATCACTTCTCTCTCTTTAGTAGTGGAGCAGCTTCGGTTTCGGGAGAAAGTCCCCAAC 691  
Db |  
QY 265 AAGGCTCGGGGAATTCGCTTACCAAAACCTTCGCGCGACGGGCAATGATAGCGGCT 324  
Db |  
QY 690 AAGGCTCGGGGAATTCGCTTACCAAAACCTTCGCGCGACGGGCAATGATAGCGGCT 631  
Db |  
QY 325 GCGCCCCACAGGACCATCGAGCGCGCGTCCAGGTTCAGGTCTTGAAGCACTTTGGG 384  
Db |  
QY 630 GCGCCCCACAGGACCATCGAGCGCGCGTCCAGGTTCAGGTCTTGAAGCACTTTGGG 571  
Db |  
QY 385 ACCGAAAGCTAAGACGGGATCCGACGCCATCTAGTTTCCCATCAACCATGTAGGCATC 444  
Db |  
QY 570 ACCGAAAGCTAAGACGGGATCCGACGCCATCTAGTTTCCCATCAACCATGTAGGCATC 511  
Db |  
QY 445 CCGCATGTAGGCGGTTGCAATGGCCAAAGTGGCGCATGTTCCAAAGTTCATCTTCA 504  
Db |  
QY 510 CCGCATGTAGGCGGTTGCAATGGCCAAAGTGGCGCATGTTCCAAAGTTCATCTTCA 451  
Db |  
QY 505 TCCGCGCACGGGATTAGTTTCAAGGGTTACCGCTCTTAAACATCTCCACGCCGACAA 564  
Db |  
QY 450 TCCGCGCACGGGATTAGTTTCAAGGGTTACCGCTCTTAAACATCTCCACGCCGACAA 391  
Db |  
QY 565 GGATAATGTGGGCTTCACTTCCAGCGCAGCGTGGCGTTCCTCCACGCCGACAA 624  
Db |  
QY 390 GGATAATGTGGGCTTCACTTCCAGCGCAGCGTGGCGTTCCTCCACGCCGACAA 331  
Db |  
QY 625 TACCTCTGTAACACGGGAGAAACCATGTGGATAGCGAATCTGGGTTGATGGCGATGGT 684  
Db |  
QY 330 TACCTCTGTAACACGGGAGAAACCATGTGGATAGCGAATCTGGGTTGATGGCGATGGT 271  
Db |  
QY 685 TAACGGGATTTCAGCAAGCGGTCAGATAGTTTCGCTTCTGCTTTCGCTTTCGAGCAAC 744  
Db |  
QY 270 TAACGGGATTTCAGCAAGCGGTCAGATAGTTTCGCTTCTGCTTTCGCTTTCGAGCAAC 211  
Db |  
QY 745 CATTTTCGCGCTCTTTCAGCAAGCGGTCAGATAGTTTCGCTTCTGCTTTCGCTTTCGAG 804  
Db |  
QY 210 CATTTTCGCGCTCTTTCAGCAAGCGGTCAGATAGTTTCGCTTCTGCTTTCGCTTTCGAG 151  
Db |  
QY 805 GCGGATACCAACTCGACCCACGCTGATGCTCGAGAGGTTTAAACGCGTGTACCGC 864  
Db |  
QY 150 GCGGATACCAACTCGACCCACGCTGATGCTCGAGAGGTTTAAACGCGTGTACCGC 91  
Db |

QY 865 CGAGCGGGAATAAGCAAGAGCGCTTAAGAGCGGCTTCGAGAGCTGCCTTCATCATGATTGA 924  
Db |  
QY 90 CGAGCGGGAATAAGCAAGAGCGCTTAAGAGCGGCTTCGAGAGCTGCCTTCATCATGATTGA 31  
Db |  
QY 925 GAGCAAGTGTCCAGTTGAATGGGTTTCAT 954  
Db |  
QY 30 GAGCAAGTGTCCAGTTGAATGGGTTTCAT 1  
Db |  
RESULT 9  
AX063771 822 bp DNA linear PAT 24-JAN-2001  
LOCUS  
DEFINITION Sequence 53 from Patent WO0100843.  
ACCESSION AX063771  
VERSION AX063771.1 GI:12541483  
KEYWORDS  
SOURCE  
ORGANISM  
Corynebacterium glutamicum  
Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
REFERENCE  
1  
Pompejus,M., Kroeger,B., Schroeder,H., Zelder,O. and Haberhauer,G.  
corynebacterium glutamicum genes encoding metabolic pathway  
proteins  
Patent: WO 0100843-A 53 04-JAN-2001;  
JOURNAL BASF AKTIENGESSELLSCHAFT (DE)  
FEATURES  
source  
Location/Qualifiers  
1..822  
/organism="Corynebacterium glutamicum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1718"  
101..802  
/note="unlabeled protein product; RXA01394"  
/codon\_start=1  
/translation="MEIFITGLIGLLSISGPNVIVIKGIRKRELLAVLLVLI  
SDVFIAGTIGLISNAAIPVLDIMRWGIAIYLLWFAVMAKDMTNKVEAPQIIE  
ETEPVVDPTPLGSAVATDTRVRVSVKQVWKPMLALVLTWLNPNAYLDA  
FVPIGVAQYGDTRWIFPAAGAFASLINFPLVGFCAALSRPLSPKVRWVNVV  
AVMTALAKLMLNG"  
CDS  
Query Match 34.6%; Score 822; DB 6; Length 822;  
Best Local Similarity 100.0%; Pred.No.6.1e-191;  
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 925 GAGCAAGTGTCCAGTTGAATGGGTTTCATGAGAGCTATATTAAACCATGTTAAGAACCA 984  
Db |  
QY 1 GAGCAAGTGTCCAGTTGAATGGGTTTCATGAGAGCTATATTAAACCATGTTAAGAACCA 60  
Db |  
QY 985 TCATTTTACTTAAGTACTTCCATAGGTTCAGATGGTATCATGAAATCTTATTACAGG 1044  
Db |  
QY 61 TCATTTTACTTAAGTACTTCCATAGGTTCAGATGGTATCATGAAATCTTATTACAGG 120  
Db |  
QY 1045 TCTGCTTTTGGGGCCAGCTTTTTCATCTGTCATCGACCGAGATGCTACTGTGATTAA 1104  
Db |  
QY 121 TCTGCTTTTGGGGCCAGCTTTTTCATCTGTCATCGACCGAGATGCTACTGTGATTAA 180  
Db |  
QY 1105 ACAGGAATTAAGCGGAGAGACTCATTTGGGTTCTTCTCGTGTGTTTAAATTTCTGAGCT 1164  
Db |  
QY 181 ACAGGAATTAAGCGGAGAGACTCATTTGGGTTCTTCTCGTGTGTTTAAATTTCTGAGCT 240  
Db |  
QY 1165 CTTTTGTGTTTATCCCGGACCTTGGCGGTTGATCTTTTGTCCAATGCCGCCGCGATCGT 1224  
Db |  
QY 241 CTTTTGTGTTTATCCCGGACCTTGGCGGTTGATCTTTTGTCCAATGCCGCCGCGATCGT 300  
Db |  
QY 1225 GCTCGATATTATGCGCTGGGGTGGCATCGCTTACCTGTTATGTTTGGCTCATGGCAGC 1284  
Db |  
QY 301 GCTCGATATTATGCGCTGGGGTGGCATCGCTTACCTGTTATGTTTGGCTCATGGCAGC 360  
Db |

1285 GAAAGACCCATGACAAACAAAGGTGGAAGCGCCACAGATCATTTGAAGAAACAGAACCAAC 1344  
 361 GAAAGACCCATGACAAACAAAGGTGGAAGCGCCACAGATCATTTGAAGAAACAGAACCAAC 420  
 1345 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1404  
 421 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 480  
 1405 GCGGTGGAAGTGAAGCGGTGATTAAGCGCGGTGGAAGCGGTGGAAGCGGTGGAAGCGGT 1464  
 481 GCGGTGGAAGTGAAGCGGTGATTAAGCGCGGTGGAAGCGGTGGAAGCGGTGGAAGCGGT 540  
 1465 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1524  
 541 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 600  
 1525 CCGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1584  
 601 CCGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 660  
 1585 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1644  
 661 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 720  
 1645 CCGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1704  
 721 CCGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 780  
 1705 CAACTGATGTTGATGGTGTAGTTTTCGCGGGTTTGGAAATC 1746  
 781 CAACTGATGTTGATGGTGTAGTTTTCGCGGGTTTGGAAATC 822

RESULT 10  
 AX244059  
 LOCUS  
 DEFINITION  
 Sequence 51 from Patent WO0166573.  
 ACCESSION  
 AX244059  
 VERSION  
 AX244059.1 GI:15859123  
 KEYWORDS  
 Corynebacterium glutamicum  
 Corynebacterium glutamicum  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O., Haberhauer, G.,  
 Kim, J. W., Lee, H. S. and Hwang, B. J.  
 Corynebacterium glutamicum genes encoding metabolic pathway  
 proteins  
 Patent: WO 0166573-A 51 13-SEP-2001;  
 BASF AKTIENGESellschaft (DE)  
 Location/Qualifiers  
 1..822  
 /organism="Corynebacterium glutamicum"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1718"  
 101..802  
 /note="unassigned protein product; RXA01394"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAC8508.1"  
 /db\_xref="GI:15859124"  
 /db\_xref="REMBASE:CA8508"  
 /translation="MEIFITGLLGSLISIQPVLVVKQIKREGLIALLVLCIL  
 SDVFLIAGTGLDGLLSNAPVILWVAVNVAANDAKMTKVAPOIIE  
 EFTPTDPTPLGSAVATDTRNVRVSVSKORVWVKMLAIVLTLNPNAYLDA  
 FVFTGGVGAQDGTGRWITPAAGAPASLIWFLVGVGAALSRPLSSPKVRWLNVV  
 AVWMTALAKMLMG"

ORIGIN  
 Query Match 34.6%; Score 822; DB 6; Length 822;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-191;

Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 925 GAGCAAGTGTCCAGTTCGAATGGGGTTCATGAGCTATATTAACCAATGTTAAGAACCA 984  
 DB 1 GAGCAAGTGTCCAGTTCGAATGGGGTTCATGAGCTATATTAACCAATGTTAAGAACCA 60  
 QY 985 TCATTTTCTAATGATCTTCATAGGTTCAGATGATCATGGAATCTTCATTTACAGG 1044  
 DB 61 TCATTTTCTAATGATCTTCATAGGTTCAGATGATCATGGAATCTTCATTTACAGG 120  
 QY 1045 TCTGCTTTTGGGGCCAGTCTTTTCTGCTTCATCGACCGCAGATGATGCTGCTGATTA 1104  
 DB 121 TCTGCTTTTGGGGCCAGTCTTTTCTGCTTCATCGACCGCAGATGATGCTGCTGATTA 180  
 QY 1105 ACAGCAATTAAGCGCGGAGATCATTCGCTTCCTTCGCTGCTGCTGCTGCTGCTGCTG 1164  
 DB 181 ACAGCAATTAAGCGCGGAGATCATTCGCTTCCTTCGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 1165 CTTTGTGTTTCAATCGCGGACCTTTGGGCGTTCATCTTTTTCCTCAATGCGCGCGATCGT 1224  
 DB 241 CTTTGTGTTTCAATCGCGGACCTTTGGGCGTTCATCTTTTTCCTCAATGCGCGCGATCGT 300  
 QY 1225 GCTCGATTTATGCTGCTGGGTGGCATCGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1284  
 DB 301 GCTCGATTTATGCTGCTGGGTGGCATCGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
 QY 1285 GAAAGACCCATGACAAACAAAGGTGGAAGCGCCACAGATCATTTGAAGAAACAGAACCAAC 1344  
 DB 361 GAAAGACCCATGACAAACAAAGGTGGAAGCGCCACAGATCATTTGAAGAAACAGAACCAAC 420  
 QY 1345 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1404  
 DB 421 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 480  
 QY 1405 GCGGTGGAAGTGAAGCGGTGATTAAGCGCGGTGGAAGCGGTGGAAGCGGTGGAAGCGGT 1464  
 DB 481 GCGGTGGAAGTGAAGCGGTGATTAAGCGCGGTGGAAGCGGTGGAAGCGGTGGAAGCGGT 540  
 QY 1465 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1524  
 DB 541 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 600  
 QY 1525 CCGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1584  
 DB 601 CCGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 660  
 QY 1585 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1644  
 DB 661 CTTGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 720  
 QY 1645 CCGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1704  
 DB 721 CCGCGCGATGACACGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 780  
 QY 1705 CAACTGATGTTGATGGTGTAGTTTTCGCGGGTTTGGAAATC 1746  
 DB 781 CAACTGATGTTGATGGTGTAGTTTTCGCGGGTTTGGAAATC 822

RESULT 11  
 AP005218/c  
 LOCUS  
 DEFINITION  
 Corynebacterium efficiens YS-314 DNA, complete genome, section 5/11.  
 ACCESSION  
 AP005218 BA000035  
 VERSION  
 AP005218.1 GI:23493016  
 KEYWORDS  
 Corynebacterium efficiens YS-314  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 Nishio, Y., Nakamura, Y., Kawarabayashi, Y., Usuda, Y., Kimura, E.,  
 1

Thu Mar 18 12:30:58 2004

Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikee, K. and  
Gojobori, T.  
Comparative Complete Genome Sequence Analysis of the Amino Acid  
Replacement Responsible for the Thermotolerance of *Corynebacterium*  
*efficiens*  
Genome Res. 13 (7), 1572-1579 (2003)  
22723752  
12840036  
2 (bases 1 to 308650)  
Kawarabayashi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and  
Director-General of Biotechnology Center.  
Submitted (17-MAY-2002) Director-General of Biotechnology Center,  
National Institute of Technology and Evaluation, Biotechnology  
Center, Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan  
[E-mail: bio@nitech.go.jp, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424]  
Kawarabayashi, Y. is officially affiliated with the National  
Institute of Advanced Industrial Science and Technology, Tsukuba,  
Ibaraki, 305-8565 Japan  
Nakamura, Y., Ikee, K., Suzuki, M. and Mashima, J. are at the  
National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan  
Ito, T. is at the Japan Biological Information Research Center,  
Koto-ku, Tokyo, 135-0064 Japan  
Yamagishi, A. is at Tokyo University of Pharmacy and Life Science,  
Hachioji, Tokyo, 192-0392 Japan  
Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co.,  
Inc., Kawasaki, Kanagawa, 210-8681 Japan  
The other authors are at the National Institute of Technology and  
Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.  
Location/Qualifiers  
1. .308650  
/organism="Corynebacterium efficiens YS-314"  
/mol\_type="genomic DNA"  
/strain="YS-314"  
/db\_xref="taxon:196164"  
complement(83..241)  
/note="CE1178"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17989.1"  
/db\_xref="GI:23493017"  
/translation="MTTGIILGTGVDGAGLKGSDIPRGVGHFRFRKPFVALQYR  
TGAGAND"  
365..742  
/note="CE1179"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17989.1"  
/db\_xref="GI:23493018"  
/translation="MEKSQLQKSTQDFASVDHLSDDARAFIDNELSRGAMHRAHLH  
LVHCACRQELARQDVTDFVTECNQNVTPSGLKAKLSLANECAPGAEADAI  
QRPEFVAVSVVRAVRQTQR"  
774..1271  
/note="CE1180, similar to AJ292256-1|CAC21724.1| percent  
identity: 26 in 155 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17990.1"  
/db\_xref="GI:23493019"  
/translation="MFSNVGGEILLFTVGLIVIGPERLPRMLDLKAAIYAARTAI  
NNAKQNLDTDFGEFDEIRKPMQIAQIRSMPTATITRALFDGDSFLDSFDPKXIM  
ADDTAGAHQHQVAKTGTTPPEAEVVERPADRIDTQSPWTPGKNDGKPKAGGGYT  
WTDLI"  
complement(1353..2480)  
/note="CE1181, similar to AL390975-15|CAC01353.1| percent  
identity: 63 in 370 aa"  
/codon\_start=1  
/transl\_table=11  
/product="putative ATP-binding protein Mrp"  
/protein\_id="BAC17991.1"

/db\_xref="GI:23493020"  
/translation="MSQVTEASRSALSRVEDPEIGKPIITELGMVKSVAIDGSDVQV  
EVLITAAQPMKTTIVSNTAAIKDIBGVQDVQVTTVMSDEQRRLRSLRGNTADP  
VIPAPQSTTVAVASGKGVKSMVTNIALAAKGLSGVILDDADIVGHVSIPGM  
LGSADPHQVDMVMPPOAHGVKVIISIAHFVEGNAPVVRGPMHRAICQFSLDFWVG  
DLITLLDIPGCTGDIATTAQLPNAELLIVTPQAAAADVAERAGSLSLQTKQKVA  
GVENMSMVPDGSSTMEVFGSGGKIAQLRSLMTGTEVFLVSGVPLDPSLRVGGDV  
GNPIASEPYSPTTAINATDLAKRKYSIVGKTLGLGVK"  
complement(2513..3271)  
/note="CE1182"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17992.1"  
/db\_xref="GI:23493021"  
/translation="MTTATPWNADAFPHVQWDLPLQLEVDHTERAVAHLOEYATDPT  
GPMYTGSLFRNLGGGDAQPVNRITRDDLSLLDVPVPGVQLLEPPEDAQSI  
DLEEVTRDYAVELPDEIQIRGLASPTDVLGASKAAHHLDAKGNLLWTAPENG  
PGRVTRSKLJARKRPKLPFIDSAIKRQNTNSGVGFENFHTVLRADKRLSHLKG  
RTEADSGNSALADLSITRVFDVVVVMADHRT"  
complement(3282..3836)  
/note="CE1183, similar to AE007003-3|AAK45527.1| percent  
identity: 57 in 159 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17993.1"  
/db\_xref="GI:23493022"  
/translation="MADYSRSLDPLVGTGRKRFKEDDDTVGAVKARVFGTQY  
LFWQITFVIMWINTGFWANWDPYFILLNASTQAAYAPLILLANRQEDRK  
STLADRRABETKADTEFLTELAVGEMVTRDYLRLHLDLDRGLERLEAKL  
DVAASRNDPHDLNAPTOGVVD"  
complement(3829..5124)  
/note="CE1184, similar to AE007003-4|AAK45528.1| percent  
identity: 52 in 417 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17994.1"  
/db\_xref="GI:23493023"  
/translation="MRGITRLVAGLACMVVRGDDTVIGRVDDVVVTRSGGQSRV  
LGLVBLVNNKRIFLPMRLIAAEIKENLISGVSLSAKFARTELOVIGDITGK  
VHTDPELQLHGRVIAEDVELEQTRTDWVIRSVAGIKRFGKRPRTLYTAPWSH  
VHGITAGAGETNQTAEILAGFEDMAADKELYGLPKQDVAGVLELDERLADIL  
QELSEDRQALLETIDERAADVLEMDPDAADLGLPEDKAGVLELMDPEEPAS  
VRLMGFNEDTVGALMTPELTMDPSTTVABALAMARNPDLPTSLASLVFVVRPTAT  
PTGKYLGCVHLQLEPPPSLVGGILDDPLPLVANDSOETAAARFATYVNLVCGPYL  
DENGHLIGAVAVDLDLMLPEWDNRDAGRPGVKEETLG"  
5264..5782  
/note="CE1185, similar to AE007003-6|AAK45530.1| percent  
identity: 60 in 155 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17995.1"  
/db\_xref="GI:23493024"  
/translation="MSMQATGSGNNRAELIKRKEGCVPGVSGFYVEAQRVDMLS  
DNEFFVAILTVGVDMLEVTGRLTWKRVLAGASGAWLGLFFGLVGLVSNLMP  
SSLITGLVMGVVPGVLAAPVYWASRKRDTTSATQIVAGRYDILCRDPRAREADMI  
ARNMGLGHPHQ"  
5805..6500  
/note="CE1186"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17996.1"  
/db\_xref="GI:23493025"  
/translation="MRTGVKALLGLVAVLGGLLVAGSVDPNRPPEHPRQVAEQK  
ENLYTAVDPDVLGHVVISLYKTVGRFAVIRHEDNLADPFDNEGSDLVIG  
CTGALLEVVPKAGLEBEYLAAVEAGTDINSGEWRDFTDGLSPSLRDLTADP  
SNAQCGPANGLFQNVPIYRNPIFSRADREILNWSAGIATTSLOELBEAKELGR  
RELKVEFVASKG"  
6660..7373

					/note="CE1187, similar to AE007262-3 [AAK65408.1] percent identity: 32 in 235 aa"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

```

FEATURES             Location/Qualifiers
     1..1771
     /organism="Corynebacterium efficiens"
     /mol_type="genomic DNA"
     /db_xref="taxon:152794"
     complement(47..919)
     /gene="lysG"
     CDS
     complement(47..919)
     /gene="lysG"
     /codon_start=1
     /transl_table=11
     /product="Lysine export transcriptional regulatory protein"
     /protein_id="BAB88826.1"
     /db_xref="GI:20065739"
     /translation="WNPHLTLLTIIDGESFNASLALSTSPSAVSQRIKALEKSVG
RVNRSQPAVATFAGEVILVQAARKMALIQASTREQLAERLDEIPLTVAINDSLSTW
FVPFAEVAGWAGVATITURKDEAHTLUSLRKGSVIGAVTREADPVAGEVILRGWR
HLPVATPELRATYVDQGVMPVILRFGNDVLQDRLEGVDGAVARRKSVSPFS
AGSGFEAVRLICGQVLLPQAAPMLAAGDVVQLDEKVDYDTPLYWQRFWELESRLAEL
TVADVDRABAGLRT"
     981..1676
     /gene="lysE"
     CDS
     981..1676
     /gene="lysE"
     /codon_start=1
     /transl_table=11
     /product="Lysine exporter protein"
     /protein_id="BAB88827.1"
     /db_xref="GI:20065740"

```

Db 897 GTGTCAGGTGGATCGGGTTTCATGAAGCAATCTTAAACATGTTCACTCTCATCATTTT 956  
 Qy 992 ACTTAAGTACTTCCATAGGTTCAGATGATCATGAATCTTCAATACAGGTCTGCTT 1051  
 Db 957 ACTTAATGTGGTTTGTGCGGACATCGGACATCGAAATCTTTGTACCGGTTTGTG 1016  
 Qy 1052 TTGGGGCCAGCTTTTACTGTGTCATCGGACCGCATGTGATGATGATTAACAGGA 1111  
 Db 1017 TTGGGACCAAGCTGCTGTTGGCCATCGGCCACAGATGCTCTGGTGATCAACAGGC 1076  
 Qy 1112 ATTAAGCGGAGGACTCAATGCGGTTCTTCTGCTGTGTTAAATCTTCACTCTTTT 1171  
 Db 1077 ATCAAGCGGAGGACATCAAGCGGTCATCATCTGCTGCTGCTCGAGGTGCTG 1136  
 Qy 1172 TTCAATGCGGACATGCGGCTGATGATCTTTTGTCAATGCGGCGGATGCTGCTGAT 1231  
 Db 1137 TTCACTCTCGGACCTCGGAGTGGGCTGATCTCGACACCGCCGATCATTTCTGAC 1196  
 Qy 1232 ATTATGCGTGGGGTGGCATGCTTACCTGTTATGTTTGGCTTCATGCGACGAAGAC 1291  
 Db 1197 ATCTGCGCTGCTGCGGATGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1256  
 Qy 1292 GCAATGACAAACAGGTGGAAGCGCCACAGATCATTTGAAGAAACAGAACCAACGTCGCC 1351  
 Db 1257 GCCCTGGGCGCCGACCGAGGTAACTTTGTC-----GAGCATTTCCGACCC 1304  
 Qy 1352 GATGACACGCTTTGGGCGGTTGCGCGGTGGCCACTGACAGCGCAACCGGGTGGGGTG 1411  
 Db 1305 GTTG---CGGCAGGCTGCGCTCCGCGCGGGGGGTGACAGCAAAACAGACCCCGGCTC 1361  
 Qy 1412 GAGTGAGCTGCTGATAGACGAGCGGTTGGGTAAAGCCCATGTTGATGCAATCTGCTG 1471  
 Db 1362 CGCATCACATCAGGACCCGCGGCTGCTGGTCAAGCCCATGCTCATGGCATTTGCTG 1421  
 Qy 1472 ACTGTGTGACCGAATGCGTMTTGAAGCGGTTTGTGTTATCGCGCGGCTGCGCGG 1531  
 Db 1422 ACTGTGCTCAATCCATGCTACCTGATGCTTCTGCTTCTGCTGCTGCTGCTG 1481  
 Qy 1532 CAATAGCGGACACCGGACCGGTGATTTTCCCGCTGCGCGGTTGCGGCAAGCTGATC 1591  
 Db 1482 CAGTACGGGAGACCGGTGCTGGATCTTCTGCGGGTCTTCCGCGCGAGCTGCTG 1541  
 Qy 1592 TGCTTCCGCTGCTGCTGCTGCGGACGATTTGACGCGGCTGCTGCGGCGCCAG 1651  
 Db 1542 TGGTTCCTCACTGGTGGTACGCGCGGCGGCTGCTGCGGCTGCTGCGGCGCGG 1601  
 Qy 1652 GTGTGGCGGTGATCAAGCTGCTGCTGCTGCTGCTGATGACCGATGCTGCTGCTG 1711  
 Db 1602 GTCTGCGCTGCTGATCAATAGTGTGCGCGCTGCTGCTGCTGCTGCTGCTG 1661  
 Qy 1712 ATGTTGATGCTTA 1725  
 Db 1662 ATCTGATGCTTA 1675

## RESULT 13

LOCUS BX248357/c 349535 bp DNA linear BCT 06-NOV-2003  
 DEFINITION Corynebacterium diphtheriae gravis NCTC13129, complete genome;  
 segment 4/8  
 VERSION BX248357 BX248353  
 KEYWORDS complete genome.  
 SOURCE Corynebacterium diphtheriae  
 ORGANISM Corynebacterium diphtheriae  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 1 (bases 1 to 349535)  
 Cordero-Tarraga, A.M., Efrat, A., Dover, J.G., Holden, M.T.G.,  
 Pallen, M., Bentley, S.D., Beera, G.S., Churcher, C., James, K.D., De  
 Zoysa, A., Chillingworth, T., Cronin, A., Dowd, L., Feltwell, T.,  
 Hamlin, N., Holroyd, S., Jagels, K., Moule, S., Quail, M.A.,  
 Rabinowitch, E., Rutherford, K., Thomson, N.R., Unwin, L.,

Whitehead, S. and Barrell, B.G. Parkhill, J.  
 The complete genome sequence and analysis of Corynebacterium  
 diphtheriae NCTC13129  
 Nucleic Acids Res. 31 (22), 6516-6523 (2003)  
 14602910  
 2 (bases 1 to 349535)  
 Cordero-Tarraga, A.M.  
 Direct Submission  
 Submitted (03-OCT-2003) Cordero-Tarraga A.M., submitted on behalf  
 of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust  
 Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:  
 amctesanger.ac.uk

## FEATURES

source  
 1. 349535  
 /organism="Corynebacterium diphtheriae"  
 /mol\_type="genomic DNA"  
 /strain="NCTC13129"  
 /db\_xref="taxon:1717"  
 /note="biotype gravis"  
 76..402  
 /locus\_tag="DIP1058"  
 76..402  
 /locus\_tag="DIP1058"  
 /note="Tsimilar to Mycobacterium tuberculosis hypothetical  
 11.0 kDa protein Rv1898 or MTCY180.20c  
 SW:1192\_MTCY180 (007734) (102 aa) fasta scores: E():  
 5.8e-16, 56.43% id in 101 aa, and to Aquifex aeolicus  
 hypothetical 17.9 kDa protein AQ\_2067 TR:067847  
 (EMBL:AE000771) (157 aa) fasta scores: E(): 4.7e-08,  
 40.19% id in 102 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Conserved hypothetical protein"  
 /protein\_id="C9E49578.1"  
 /db\_xref="GI:38199313"  
 /translation="MLIDMLLAFSVAPETPNNAEMADVAEALRVRESGLPNETN  
 AMFLIIGENDEVMVKKATDAVLAVSPRASLVKADIRPGTIGQVQKVASVERVL  
 AGDSSES"  
 88..372  
 /locus\_tag="DIP1058"  
 /note="HMMpfam hit to PF01910, Protein of unknown function  
 DUF77"  
 complement (405..4762)  
 /note="Putative iron uptake system. Low GC content  
 (58.02%)"  
 complement (405..1226)  
 /locus\_tag="DIP1059"  
 complement (405..1226)  
 /locus\_tag="DIP1059"  
 /note="Tsimilar to Streptomyces coelicolor putative  
 iron-siderophore uptake system ATP-binding component  
 SC151\_25C TR:Q98215 (EMBL:AL109848) (301 aa) fasta scores:  
 E(): 1.9e-42, 50.58% id in 255 aa, and to Escherichia coli  
 ferric enterobactin transport ATP-binding protein fepC  
 b0588 SW:FEPC\_ECOLI (P23878) (271 aa) fasta scores: E():  
 1.1e-41, 49.61% id in 262 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Putative iron-siderophore uptake system  
 ATP-binding component"  
 /protein\_id="C9E49580.1"  
 /db\_xref="GI:38199314"  
 /translation="MTSTSPILKTSVTLANDKRTSTLTVEVQCGQTALIGNGC  
 GKSTLLKSFARILQPHTCQVTLGNDPDLSTHTKTAQRIALLQPTALAPALIVELV  
 GSGFPYCTLLRQWSTDAQVDSALLATGAQLRRHRTVELSGGQGVNLANLAQ  
 NTPVVLDEPTTLDVAHQQLLELALRQQLGRTVTVLVLDLQCAVRVADHLIVMK  
 CGAVYAGCAPSSIMTPQLISEVFDINVDVHTIDHVVIVKTPMPQDPVSVLENLQ"  
 complement (564..1133)  
 /locus\_tag="DIP1059"  
 /note="HMMsmart hit to SM00382, ATPases associated with a  
 variety of cellular activities"  
 complement (567..1130)  
 /locus\_tag="DIP1059"

## misc\_feature

## misc\_feature

Thu Mar 18 12:30:58 2004

```

/misc_feature
/note="HMPfam hit to PF00005, ABC transporter"
/complement(579..797)
/locus_tag="DIP1059"
/note="ProfileScan hit to PS0100, 2nd half motif for
nucleotide binding, associated with p-loop."
/complement(753..797)
/locus_tag="DIP1059"
/note="ScanRegExp hit to PS00211, ABC transporters family
signature."
/complement(1047..1124)
/misc_feature
/locus_tag="DIP1059"
/note="ProfileScan hit to PS0101, P-loop nucleotide
binding motif (does not find all)."
/complement(1086..1109)
/misc_feature
/locus_tag="DIP1059"
/note="ScanRegExp hit to PS00017, ATP/GTP-binding site
motif A (P-loop)."
/complement(1223..2182)
/gene
/locus_tag="DIP1060"
/CDS
/locus_tag="DIP1060"
/note="Similar to Streptomyces coelicolor putative
iron-siderophore uptake system transmembrane component
sc151_26C TR:Q9S214 (EMBL:AL109848) (375 aa) fasta scores:
E(): 5e-40, 40.74% id in 324 aa, and to Escherichia coli
ferric enterobactin transport system permease protein fepg
fepg or B0589 SW:FEPG ECOLI (P23877) (330 aa) fasta
scores: E(): 1.8e-45, 44.09% id in 322 aa"
/transl_table=1
/transl_table=11
/product="Putative iron-siderophore uptake system
transmembrane component"
/db_xref="GI:38199915"
/translations="MRALVLSLLVTAAYSLVPGAGLSTWQLTSTPLAHTVM
QWLRPVVTGLVGAALATAGSLTRNPLGSPDIIIGFTGAYTVIAAFLGWG
FGATVIGALVGLAVSVVIALSVTRIDGLRIILVGLIGISALSNLWLRIGELD
TALSAASWAGSLNGLRWLTATPACCVLAIIITPIRLRELDVLSIGDDLVAGLRL
LQUTKLLLVGVPLVAITAVAGPIAFVALASPHARALTSNRPVETSVIGALL
IIVADLIGRLFPYPTQLPVGLVTVIGGYMLLARGSKENS"
/complement(2108..2182)
/misc_feature
/locus_tag="DIP1060"
/note="Signal peptide predicted for DIP1060 by SignalP 2.0
HMM (Signal peptide probability 0.966) with cleavage site
probability 0.180 between residues 25 and 26;
signal-peptide site"
/complement(1244..2116)
/misc_feature
/locus_tag="DIP1060"
/note="HMPfam hit to PF01032, FeCCD transport family"
/order(complement(2105..2161), complement(1979..2044),
complement(1893..1939), complement(1805..1870),
complement(1718..1783), complement(1607..1657),
complement(1445..1510), complement(1319..1384),
complement(1247..1303))
/locus_tag="DIP1060"
/note="9 probable transmembrane helices predicted for
DIP1060 by TMHMM2.0"
/complement(1250..1564)
/misc_feature
/locus_tag="DIP1060"
/note="BlastProDom hit to PD001557, PD001557"
/complement(2184..3191)
/gene
/locus_tag="DIP1061"
/complement(2184..3191)
/CDS
/locus_tag="DIP1061"
/note="Similar to Streptomyces coelicolor putative
iron-siderophore uptake system transmembrane component
sc151_27C TR:Q9S213 (EMBL:AL109848) (348 aa) fasta scores:
E(): 8e-36, 36.53% id in 323 aa, and to Escherichia coli
ferric enterobactin transport system permease protein fepd
or B0590 SW:FEPD ECOLI (P23876) (334 aa) fasta scores:
E(): 5.2e-40, 40.37% id in 322 aa"
/codon_start=1
/transl_table=11
/product="Putative iron-siderophore uptake system
transmembrane component"
/db_xref="GI:38199916"
/translations="MLQPRVRLPLVITFLVILLCTIIMSLGVGARNTSPDILA
ALRHGDPNNSHIWGRIPPTLATAAGASISLAGVLIOALTRNPLADTVFGINAG
AAFTVIGTALGASLSHISIPFALLAGLACVVALSWNSCKGSDPLRLVLAVGAL
SAIITGTGDSLSVNPQAFDLKSNMWNIDAGSVQPNVAGVGLILGLIATVTCMQ
INALSLDELAITMGASAKTRLUFTATVAVASATAAGVITFLGLIMVPHARWIV
GNLLRLVATSLIGIITVLSADILGRITVPGEPPAGVVAFTGAPFLIAYATKRKE
1"
/misc_feature
/complement(3090..3191)
/locus_tag="DIP1061"
/note="Signal peptide predicted for DIP1061 by SignalP 2.0
HMM (Signal peptide probability 1.000) with cleavage site
probability 0.663 between residues 34 and 35;
signal-peptide site"
/complement(2199..2525)
/misc_feature
/locus_tag="DIP1061"
/note="BlastProDom hit to PD001557, PD001557"
/complement(2205..3095)
/misc_feature
/locus_tag="DIP1061"
/note="HMPfam hit to PF01032, FeCCD transport family"
/order(complement(3096..3161), complement(2934..2990),
complement(2838..2903), complement(2766..2831),
complement(2670..2735), complement(2544..2609),
complement(2394..2459), complement(2289..2354),
complement(2211..2276))
/locus_tag="DIP1061"
/note="9 probable transmembrane helices predicted for
DIP1061 by TMHMM2.0"
/3239..4282
/locus_tag="DIP1062"
/3239..4282
/CDS
/locus_tag="DIP1062"
/note="Similar to Escherichia coli iron(III)
dicarboxylate-binding periplasmic protein precursor FecB or
B4290 SW:FECB ECOLI (F15028) (300 aa) fasta scores: E():
Query Match 30.5%; Score 724.4; DB 1; Length 349535;
Best Local Similarity 58.2%; Pred. No. 1.2e-166;
Matches 1551; Conservative 0; Mismatches 931; Indels 38; Gaps 3;
QY 81 ACTAAGCCGCAATCCCTCGATCTCTGTCATCAACGAGGCTCTGTGAGCTTAGCTAGAG 140
DB 38257 ACAGTGGATGTAGTCCCTCGAGGCTGCATCCACTACCGCTTGAGTCAACCGCTTAGGG 38198
QY 141 ATCTAGATTCACGCGCCATCGTTCCTCAATACATCGGTGTGTCAATGGGTATCTCATCGA 200
DB 38197 CTGGAGATTCACCGCCATCGCTGCCAATACAGCGGAACCTTCGTGCACCTCGCTATCTA 38138
QY 201 GGAGGATCACTCTCTCTGTTTATAGCATGGAGAGCTGGTGGTTTCGGGAGCAAGTCCCC 260
DB 38137 GCATCACTACTCAACAGAAATCCAAAAGCGGTAAATATATCGGGAGCATTCGCC 38078
QY 261 AACCAAGGCTCGGCAATTCCTCAACAAACCTTCGCCGAGCGGCAATCGGATAGCG 320
DB 38077 AGCCCAATCCCAACGCGCGCTTCATGATGATCCCTCCGATGAGGATCATTAACAC 38018
QY 321 GCTCGGCCCCACAGGACCATCGAGCGCGCCGCTCCATGATGATCCCTCCGATGAGGATCAT 380
DB 38017 GTCGCCCTGCGGAGCTTCCAGAGACGCTTAATATCATCGGTCTTGAAGCCATCCC 37958
QY 381 TGCGACCAAGGCTAGAGCGGCTATCGGACCCATCTAGTTTCCCATCAACCATGTAGG 440
DB 37957 GAGGCGCAATTCGCAATCTAGGCTCCGCTCCCAATCAACAGGACCATCTATGGTGTAGC 37898
QY 441 CATCCGCAATGAGGCGGTTCATAGGCGCAAGTGGCGCATGTTTCCAGTTCTTACTACTT 500
DB 37897 GATCCCAAGGCTCGGACTTGCCTGCAAAATATCGGAATGCACCTTAACGCTATGAGT 37838
QY 501 CACATCCGCGCCACGGGATAGTTTCAACGGGTTACCGCTCTCTAAACATCTCCACGCCGA 560

```



Db 37837 CGCAACAGATACAGGTGCTCTTTCCCGCCGTACAGCGCAAGGACATCTCCGCGACGCA 37778  
 Qy 561 GCAAGATATCTGTGGCTTATCTTCCAGCGAGCGTGGGCTGGCTCCACCCCAAG 620  
 Db 37777 ACAAGCAACGAGTGGATCTGCTCTGATACGATATGCAATGTGGGCTTATCATCT 37718  
 Qy 621 AAGTACCTCGTTGAACACGGGAGGAACCAATGTGATAGCGAATCTGGTGTGATGGGA 680  
 Db 37717 CCGCCACACTTCCCATCACCGCGGAAACCAATGTGATAGCGAATCTGGTGTGATGGGA 37658  
 Qy 681 TGGTTAAGCGGATTTGACGAAGCGCTCCAGATAGTGGCTTATGTTCTGCTTGCAGCA 740  
 Db 37657 CAGACATGGAATCGCTCAATCTCTCCCGCACTGAGCGTTTGTTCGGCTGCGAA 37598  
 Qy 741 ACACATTTTCGGCTGCTGTGACAGGACTTCCACCGCTCGTGTGCTTGGCGGTT 800  
 Db 37597 GTGCCATAGCTGCGCGGATTCACACGATCTCCGCGGATCAGTTGCGAGTGAUTGGT 37538  
 Qy 801 GGGTGGCGGATACCAACACTCGACCCAGCTGATGCTCGAGAGCTTTAAAGCGCTGACTCA 860  
 Db 37537 ACGTACGGGCTACACTACTCGACAGTCTCCCGCTCCAAAGCTTAACTCTGACTTA 37478  
 Qy 861 CGCGGAGGGGAATGGAAGGGCTAAGAGGGCTTCCGAGCTGCTTCCATCATATGA 920  
 Db 37477 CTGCTGAAGGAGAAATTCCTAATGCAAAAGCGGATCTTCGAGCGCTGCTGCGATTA 37418  
 Qy 921 TTGAGACAAAGTGTCCAGTGTGAATGGGCTTCATGAGCTATATTAACCATGTTAAGAA 980  
 Db 37417 TTGCGAGCAAGTGTCTAATGAAGCGGATTCATGAAGCAATTCATCAACCACTGAAAG 37358  
 Qy 981 CCAATCATTTTACTTACTTCCATAGCTCAGATGCTGATCATGGAATCTTCAATTA 1040  
 Db 37357 TCTGTATTTTACTTACTTCCATAGCTCAGATGCTGATCATGGAATCTTCAATTA 37298  
 Qy 1041 CAGTCTGCTTTTGGGGCCAGTCTTTTACTGCTCATCGACCGCGAATGACTGTTGA 1100  
 Db 37297 CTGGAATCTCTAGGCTCTGCTCTGATCGTGGCATCTGGGCGCAAAATGGCTGATTA 37238  
 Qy 1101 TTAAACAAGGAATTAAGCGGAGGACTCAATGGCTTCTCTGCTGTTGTTAAATTTCTG 1160  
 Db 37237 TTGCTGAAGGATTAAGCGGAGGCTCAATCCCAATCTGCTGTTGATGACTCTCTG 37178  
 Qy 1161 ACGTCTTTTGTGATCGCGGCACTTGGGGCTTATCTTTGCTGATGCTGCTGCTGCTGCT 1220  
 Db 37177 ACGTGAATCTCTATTTTGGAGGCACTGGCGGTGTGGCGCAATTAAGTCAATGCTGAG 37118  
 Qy 1221 TCGTCTCGATATTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1280  
 Db 37117 TCGCACTAGTAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37058  
 Qy 1281 CAGCGAAGACCGATGACAAACAGGTGGAAGCGCCAGATCATTTGAAGAAACAGAAC 1340  
 Db 37057 GTTTAAGGAAGCAATTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37001  
 Qy 1341 CAACGCTGCCGATGACACGCTTTTGGCGGTTCGGCGGTGCGCACTGACACGCGCAACC 1400  
 Db 37000 CGGTGGCTACGAAACAGTGGCTGACGCTATCATCCGAGTGAATCACTAAACACGCAAA 36941  
 Qy 1401 GGGTGGGTGAGGTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1460  
 Db 36940 AAGCTGAGCCCAAG-----TCAGCGCGGCGAGCTGGGTGAACCGAGCTGTCAG 36890  
 Qy 1461 CAATCGTGTGACTGCTGTTGAACCCGATTTTGGACGCTGTTTGGACGCTGTTTATCGCG 1520  
 Db 36890 CGTTGGCTTTTCAAGTGGCTTAACTCCAGCTGCTTATAGACGCTTCTTGTATGTTGGAG 36830  
 Qy 1521 GGTTCGGCGGCAATGAGCGGACACGCGAGGTGATTTTCGGCTGCGGCTGCTGCGG 1580  
 Db 36829 GAATCGCCACCAACGCGTCCGAGCGGCGCTGGGTCTTCGGCTTGGGCTGCTGCTG 36770  
 Qy 1581 CAAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1640  
 Db 36769 CGAGTCTCAGTGGTTCGCTTATTTGATATACCTCCACGCTTTTCTACCGCTGCTAT 36710

Qy 1641 CCAGCCCAAGGTGTGGCTGTGGATCAACGCTGCTGTGGCAGTTGTGATGACCGCATGG 1700  
 Db 36709 CGCGGCCAGCGTGTGGGATATATCAACATAGATCGAATCATCATGATGATCATGT 36650  
 Qy 1701 CCAATCAATGATGTGATGG-----GTTAGTTTTCGG 1734  
 Db 36649 GCGACGCTCTATCATGCACTAGGCTCATTTAGGTACCGCGCTTTAATCTGT 36590  
 Qy 1735 GGTTTGGAAATGGTGGCTTCCGCCAAATTTGATGCGCGCTGTGGGAAATCTCATC 1794  
 Db 36589 AACCTTCGAGGCGTAGCACCGCCCAATTAATTAATTCACGCTCTTGGCAACAGTGT 36530  
 Qy 1795 GATCGCTTCAACTCGCGCTGAGAAATCTCAAGTGTGTGATGATCAAGCGCTGTGTC 1854  
 Db 36529 AATATAATTTGGCTTCTCAACGCTGATCTAGATTAATTAAGCTTCCAGATTTTGTATC 36470  
 Qy 1855 CAGTGTCTCAATGACGAGCACCAATCAATGACCTGCTACGATATCCGCGCTGCTACT 1914  
 Db 36469 CAATCTGCTTACCGACGATCTCCGATCAACGATCTGGTTACCGTTTGGCGCCATATC 36410  
 Qy 1915 TCCCTTGTCTCGCGACGACCCATGCAAGCGCATCTGCGCAAGTCACTTGGCGCTTCTG 1974  
 Db 36409 GCTTTGTTCCCGAGACACCAAGATAGCTGTGCGCAAGTTCGCGACGCTGTAA 36350  
 Qy 1975 GCGATGTCAATGAGCTTGGGACCAATCAATTAATTTGCTACGTTCAATGCTCCCTAGA 2034  
 Db 36349 CGCTAAATCAATTCAGTGCACGACCATGTCAGATTTTTCAGCTTTGAGCATCTCTTGGC 36290  
 Qy 2035 CAGGAGCTTACCTGCTGGCGGGGAACTCTTGGAAATTCATCGAGATATTTGCTGCT 2094  
 Db 36289 CAGGACTTCCCTGCTGCTGCGGTGATCCACTGCGACGCTCGAGATACCTGCTCAGT 36230  
 Qy 2095 GAGCGGCTTCCGAGTGTGAGAAAGCAATGAGCGCAAGCAATTTGCGAGCTGA 2154  
 Db 36229 TAAAGCGCTTGTGCAAGTGTGCAAAATGCAATAAACCCAGTCCGTTATTCGCGCGA 36170  
 Qy 2155 GTGCAACAAAGTCTCAGCTGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2214  
 Db 36169 CTCTAACAGGTTCTCGCATCTTCACTGCTGCTTCTTCTACCGCGGTTCAAGATAGAATA 36110  
 Qy 2215 GCTTGTGCTGATGATGAGAGGCGGCGCTCTCTCGCGATGATCACTGCGCTCGC 2274  
 Db 36109 GCTCGTGTGATGATGAGAGGCGGCGCTCTCAACAGTGAATTTCCAGATCTCTGC 36050  
 Qy 2275 TGTGAGCTTGAAGCTGAGAGAAATACCAAGTAAAGCTTTTCCAGACGCAACAT 2334  
 Db 36049 GGTTAGTTCAGGCTCCATAGAGAAATTTCCACATAGCGAGCTTTCGCGATGCCAGAT 35990  
 Qy 2335 GTCAAGCAATGCTGATGCTGCTTCTTCCAAAGGAGTATCT 2374  
 Db 35989 GTCTCGCAAGCATATGCAATTTCTTCCAAAGGCTATCT 35950

## RESULT 14

AX643028  
 LOCUS AX643028 711 bp DNA linear PAT 24-FEB-2003  
 DEFINITION Sequence 7 from Patent EP1266966.  
 ACCESSION AX643028  
 VERSION AX643028.1 GI:28550158  
 KEYWORDS  
 SOURCE Corynebacterium glutamicum  
 ORGANISM Corynebacterium glutamicum  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacteriineae; Corynebacteriaceae; Corynebacterium.  
 REFERENCE 1  
 AUTHORS Gunji, Y. and Yasueda, H.  
 TITLE Method for producing L-lysine or L-arginine by using methanol  
 JOURNAL assimilating bacterium  
 JOURNAL Patent: EP 1266966-A 7 18-DEC-2002;  
 FEATURES Ajinomoto Co., Inc. (JP)  
 location/Qualifiers  
 source 1. .711





us-09-105-117k-1.rge

Thu Mar 18 12:30:58 2004

Search completed: March 15, 2004, 22:01:55  
Job time : 9515 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 15, 2004, 10:09:44 ; Search time 968.046 Seconds  
(without alignments)  
10418.122 Million cell updates/sec

Title: US-09-105-117K-1  
Perfect score: 2374  
Sequence: 1 ccatttgcgaagtgttac.....ttcttcacaaggaggtatct 2374

Scoring table: IDENTITY NUC  
Gapop 10\_0, Gapext 1.0

Searched: 337863 seqs, 2124099041 residues  
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: N\_Geneseq\_29Jan04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2003as.\*  
8: Geneseqn2003bs.\*  
9: Geneseqn2003cs.\*  
10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2374	100.0	2374	2	AAT96816 DNA encod
2	2374	100.0	2374	9	ADB66196 DNA fragm
3	2374	100.0	349980	5	Aah68528 C glutami
4	993	41.8	993	4	Aaf71777 Coryneb
5	993	41.8	993	4	Aas96096 C. glutam
6	870	36.6	870	5	Aah68421 C. glutam
7	822	34.6	822	4	Aaf71779 Coryneb
8	822	34.6	822	4	Aas96098 C. glutam
9	711	29.9	711	8	ACC80941 LysE prot
10	708	29.8	708	5	Aah68420 C glutami
11	698.6	29.4	1568	4	Aah45375 C. thermo
12	696.8	29.4	627	5	Aah68419 C glutami
13	627	26.4	627	5	Aaf68077 Coryneb
14	485.2	20.4	993	4	ACA29651 Prokaryot
15	305.8	12.9	1095	7	ACA29651 Prokaryot
16	302.8	12.8	879	7	ACA25567 Prokaryot
17	132.6	5.6	1041	7	ACA25567 Prokaryot
18	123.8	5.2	897	7	ACA26879 Prokaryot
19	122.8	5.2	909	7	ACA38559 Prokaryot
20	122.8	5.2	15239	2	AAT33536 BCG delet
21	121.2	5.1	912	7	ACA40574 Prokaryot
22	121.2	5.1	110000	4	Continuation (23 o
23	121.2	5.1	110000	4	Continuation (23 o

Aca24017 Prokaryot  
Aca26520 Prokaryot  
Aca32095 Prokaryot  
Add13356 C. glutam  
Aca19150 Prokaryot  
Aca51519 Prokaryot  
Aca50894 Prokaryot  
Aaf71729 Coryneb  
Aad55755 Escherich  
Aca45509 Prokaryot  
Abk72786 Bacillus  
Abq81844 Bifidobac  
Aas52558 E. coli D  
Aca24884 Prokaryot  
Aca53702 Prokaryot  
Aas59552 Propionib  
Acf64481 Propionib  
Aca35500 Prokaryot  
Aah66360 C glutami  
Acf71729 Photornab  
Continuation (50 o

ALIGNMENTS

RESULT 1  
ID AAT96816 standard; DNA; 2374 BP.  
XX  
AC AAT96816;  
XX  
DT 12-MAR-1998 (first entry)  
XX  
DE DNA encoding LysG, LysE and ORF3 from Corynebacterium glutamicum.  
XX  
KW LysG; LysE; ORF3; lysine transport; regulatory protein; export protein;  
Microbial production; amino acid; animal feed additive; ds.  
XX  
OS Corynebacterium glutamicum.  
XX  
FH Key Location/Qualifiers  
FT CDS complement(82..954)  
FT /tag= a  
FT /label= LysG  
FT /tag= b  
FT /label= LysE  
FT complement(1723..2373)  
FT /tag= c  
FT /label= orf3  
XX  
DE19548222-N1.  
XX  
26-JUN-1997.  
XX  
22-DEC-1995; 95DE-01048222.  
XX  
22-DEC-1995; 95DE-01048222.  
XX  
(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.  
XX  
Vrljic M, Eggeling L, Sahn H;  
XX  
WPI, 1997-333867/31.  
XX  
P-PSDB; AAW37714, AAW37715, AAW37716.  
XX  
Increasing microbial production of amino acids, especially lysine - by  
improving export carrier activity or corresponding gene expression, also  
new export and regulatory genes from Corynebacterium.  
XX  
Claim 23 and 26; Page: 16pp; German.  
PS



Db 1861 CTCACCTGACGAGCAACCAATCAATGCACTGGTACGATACGCGCGGTACTCTCTCTG 1920  
 Qy 1921 CTCGCGAGACCCATGCAAGCGCCATCTGCGCAAGTGAATGCGCGCGGTCTCTGGCGAT 1980  
 Db 1921 CTCGCGAGACCCATGCAAGCGCCATCTGCGCAAGTGAATGCGCGCGGTCTCTGGCGAT 1980  
 Qy 1981 GTCAATTGAGCTTGGCGACCAATATCAATATGTTTCACTGTTCAATGCGCGGTCTCTGGCGAT 2040  
 Db 1981 GTCAATTGAGCTTGGCGACCAATATCAATATGTTTCACTGTTCAATGCGCGGTCTCTGGCGAT 2040  
 Qy 2041 CTTACCTGCTGGCGCGGAGACCCCTCTGGAATTCATCGAGATATGTTGCGTGAGCAG 2100  
 Db 2041 CTTACCTGCTGGCGCGGAGACCCCTCTGGAATTCATCGAGATATGTTGCGTGAGCAG 2100  
 Qy 2101 GCGCTGCGCAAGTGGTGAGAGCAATGACGCCAAGACCAATGTTGGCGAGCTGACTGCA 2160  
 Db 2101 GCGCTGCGCAAGTGGTGAGAGCAATGACGCCAAGACCAATGTTGGCGAGCTGACTGCA 2160  
 Qy 2161 CAAGTTCTCACCGTCACTGCGCGGTCTCTGCAAGTATGATGATGATGATGATGATGATG 2220  
 Db 2161 CAAGTTCTCACCGTCACTGCGCGGTCTCTGCAAGTATGATGATGATGATGATGATGATG 2220  
 Qy 2221 CTGATGATCAAGAGCGGCGCGGTCTCTGCAAGTATGATGATGATGATGATGATGATGATG 2280  
 Db 2221 CTGATGATCAAGAGCGGCGCGGTCTCTGCAAGTATGATGATGATGATGATGATGATGATG 2280  
 Qy 2281 CTCTGACCGTGAAGAGCAATATCCACGTAAGAGCGCTTCCAGAGCAATGATGATGATG 2340  
 Db 2281 CTCTGACCGTGAAGAGCAATATCCACGTAAGAGCGCTTCCAGAGCAATGATGATGATG 2340  
 Qy 2341 CAATGGGTACATGGTCTTCTTCCAAAGGAGTATCT 2374  
 Db 2341 CAATGGGTACATGGTCTTCTTCCAAAGGAGTATCT 2374

RESULT 2  
 ADB66196  
 ID ADB66196 standard; DNA; 2374 BP.  
 XX ADB66196;  
 XX 04-DEC-2003 (first entry)  
 XX DNA fragment containing C. glutamicum LysE and LysE genes.  
 XX L-arginine production; coryneform bacteria; lysE; arginine repressor;  
 KW argK; liver function promoting agent; amino acid infusion;  
 KW amino acid pharmaceutical; LysE; ds.  
 XX Corynebacterium glutamicum.  
 XX Key Location/Qualifiers  
 FH 1025..1726  
 FT /\*tag= a  
 FT /product= "Protein encoded by LysE gene"  
 FT  
 XX US2003113899-A1.  
 XX 19-JUN-2003.  
 XX 17-JUL-2002; 2002US-00196232.  
 XX 25-JUL-2001; 2001JP-00224586.  
 XX (AJIN ) AJINOMOTO CO INC.  
 XX Yamaguchi M, Ito H, Gunji Y, Yasueda H;  
 XX WPI; 2003-708853/67.  
 DR P-PSDB; ADB66197.  
 XX A microorganism comprising enhanced expression of the lysE gene is useful  
 PT for enhanced production of L-arginine.  
 PT

XX Example 4; Page 26-27; 36pp; English.  
 PS The present invention relates to a method for producing L-arginine in a  
 CC microorganism (e.g. coryneform bacteria) that has L-arginine producing  
 CC ability and has been modified for enhanced expression of the lysE gene.  
 CC The microorganism is also modified so that an arginine repressor (argR)  
 CC does not function normally. The method of the invention is useful for the  
 CC enhanced production of L-arginine which is useful in liver function  
 CC promoting agents, amino acid infusion and comprehensive amino acid  
 CC pharmaceuticals. The present sequence represents a DNA fragment  
 CC containing Corynebacterium glutamicum LysE and LysE genes. Note: The  
 CC present sequence is given as SEQ ID No:24 in the sequence listing but is  
 CC referred to as SEQ ID No:25 in the rest of the specification.  
 XX Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 2374; DB 9; Length 2374;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 COATTTGCTGAAGTGTACTCTGCTGCGCCCAATTCCTGCGCGCAAGAGTGAAGAAC 60  
 Db 1 COATTTGCTGAAGTGTACTCTGCTGCGCCCAATTCCTGCGCGCAAGAGTGAAGAAC 60  
 Qy 61 CTTGAACCTTTTTCAGAGTAACTAAGCGCGCAATCCCTCGATTGTGTCATCAACGACGG 120  
 Db 61 CTTGAACCTTTTTCAGAGTAACTAAGCGCGCAATCCCTCGATTGTGTCATCAACGACGG 120  
 Qy 121 GTCTGTGAGTCTAGCTAGAGATCTAGATTCCAGCGCGCATCGTTGCAATCATCTGGTGT 180  
 Db 121 GTCTGTGAGTCTAGCTAGAGATCTAGATTCCAGCGCGCATCGTTGCAATCATCTGGTGT 180  
 Qy 181 GTCAATGGGTATCTCATCGAGAGGATCACTTCTCTCTCTTTTAGCATGGAGCAGCTTG 240  
 Db 181 GTCAATGGGTATCTCATCGAGAGGATCACTTCTCTCTCTTTTAGCATGGAGCAGCTTG 240  
 Qy 241 GGTTCGCGAAGTGTCTCCCAACCGAGGCTCGCGCAATTCCTCAACAAACCTTCGCG 300  
 Db 241 GGTTCGCGAAGTGTCTCCCAACCGAGGCTCGCGCAATTCCTCAACAAACCTTCGCG 300  
 Qy 301 CGAGCGGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
 Db 301 CGAGCGGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
 Qy 361 ACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
 Db 361 ACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
 Qy 421 TTTTCCATCAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
 Db 421 TTTTCCATCAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
 Qy 481 GGTTCGCGAAGTGTCTCCCAACCGAGGCTCGCGCAATTCCTCAACAAACCTTCGCG 540  
 Db 481 GGTTCGCGAAGTGTCTCCCAACCGAGGCTCGCGCAATTCCTCAACAAACCTTCGCG 540  
 Qy 541 TAAACATCTCCAGCGCGAGCAAGATATGATGATGATGATGATGATGATGATGATGATGATG 600  
 Db 541 TAAACATCTCCAGCGCGAGCAAGATATGATGATGATGATGATGATGATGATGATGATGATG 600  
 Qy 601 GAGCGTGTCTCCAGCGCGAGCAAGATATGATGATGATGATGATGATGATGATGATGATGATG 660  
 Db 601 GAGCGTGTCTCCAGCGCGAGCAAGATATGATGATGATGATGATGATGATGATGATGATGATG 660  
 Qy 661 CGAATCTGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
 Db 661 CGAATCTGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
 Qy 721 TTTAGTTTCTGCTGCGAGCAACCATATTTTCCGCGCTGCTTGCACAGGAGCTTCCACCGCG 780  
 Db 721 TTTAGTTTCTGCTGCGAGCAACCATATTTTCCGCGCTGCTTGCACAGGAGCTTCCACCGCG 780



XX The present invention provides a number of nucleotide and protein  
CC sequences from the Corynebacterium glutamicum. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of corynebacterium, measuring expression amount and analysing  
CC the expression profile or expression pattern of a gene derived from  
CC Corynebacterium, and identifying a homologue of a gene derived from  
CC Corynebacterium. Corynebacterium bacteria are useful for producing amino  
CC acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a nucleic acid described  
CC in the exemplification of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the European Patent Office

XX SQ Sequence 349980 BP; 80727 A; 91049 C; 97618 G; 80586 T; 0 U; 0 Other;  
Query Match: 100.0%; Score 2374; DB 5; Length 349980;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATTTGCTGAAGGTGTTACTCTGCTGGCCCAATTCCTGGCGGCGAAGTGAATAAC 60  
DB 129968 CCAATTTGCTGAAGGTGTTACTCTGCTGGCCCAATTCCTGGCGGCGAAGTGAATAAC 129909

QY 61 CCTGAACCTTTTCAGAAAGTAACTAAGCGCGCAATCCCTGATGCTGATCAACAGCGC 120  
DB 129908 CCTGAACCTTTTCAGAAAGTAACTAAGCGCGCAATCCCTGATGCTGATCAACAGCGC 129849

QY 121 GTCTGTGAGCTAGTAGAGATCTAGATTCAGAGCGCATTCGTTGCCAATACATCGGTGT 180  
DB 129848 GTCTGTGAGCTAGTAGAGATCTAGATTCAGAGCGCATTCGTTGCCAATACATCGGTGT 129789

QY 181 GTCAATGGGTATCTCATCGAGAGGATCACTTCTCTGCTTTTAGCATGGAGCAGCTTG 240  
DB 129788 GTCAATGGGTATCTCATCGAGAGGATCACTTCTCTGCTTTTAGCATGGAGCAGCTTG 129729

QY 241 GGTTCGGGAAGAGTCCCAACCAAGGCTCGGGAATTCCTCACCAGAACTCTCCG 300  
DB 129728 GGTTCGGGAAGAGTCCCAACCAAGGCTCGGGAATTCCTCACCAGAACTCTCCG 129669

QY 301 CGACGGACATAGATAGCGCTCGCCGCCACAGAGCCATCGACGCGCCGCTCCAGTC 360  
DB 129668 CGACGGACATAGATAGCGCTCGCCGCCACAGAGCCATCGACGCGCCGCTCCAGTC 129609

QY 361 AGCGTCTTGAAGCACATCTTTGGACCGAAGGCTAAGACGGCATCGACGCCAATCTAG 420  
DB 129608 AGCGTCTTGAAGCACATCTTTGGACCGAAGGCTAAGACGGCATCGACGCCAATCTAG 129549

QY 421 TTTCCCATCAACCATGTAGGCATCCCGCAATGAGGGGTTCGAATGSCCAAGTGGCGCAT 480  
DB 129548 TTTCCCATCAACCATGTAGGCATCCCGCAATGAGGGGTTCGAATGSCCAAGTGGCGCAT 129489

QY 481 GGTTCAGAGTTCTACTACTTACATCCGCGCACAGGATTTAGCTTACGGGTTACCGCTCC 540  
DB 129488 GGTTCAGAGTTCTACTACTTACATCCGCGCACAGGATTTAGCTTACGGGTTACCGCTCC 129429

QY 541 TAAACATCTCCACCGCGCAGCAAGATAATGTGCGCTTCATCTTCCAGCGCAGGT 600  
DB 129428 TAAACATCTCCACCGCGCAGCAAGATAATGTGCGCTTCATCTTCCAGCGCAGGT 129369

QY 601 GAGCGTGTCTCCACCCAGAGCTTACCTGTTGACACGGGAGAAACCATGTGATAG 660  
DB 129368 GAGCGTGTCTCCACCCAGAGCTTACCTGTTGACACGGGAGAAACCATGTGATAG 129309

QY 661 CGAATCTGCGTGTGATGCGATGTTAAACGGATTTAGCAAGGCGCTCCAGATAGTGGC 720  
DB 129308 CGAATCTGCGTGTGATGCGATGTTAAACGGATTTAGCAAGGCGCTCCAGATAGTGGC 129249

QY 721 TTTAGTTTCTGCTGACAGCAACACCATTTTCGCGCTGTGTCAGAGACTTACCGCG 780  
DB 129248 TTTAGTTTCTGCTGACAGCAACACCATTTTCGCGCTGTGTCAGAGACTTACCGCG 129189

QY 781 TTCGGTGTCTTGGCCGCTGGGTGGCGGATACCACTCGACCCACGCTGATGCTCGAG 840

DB 129188 TTCCGCTGCTTTTGGCGCGTGGTGGCGGATACCAACTCGACCCACGATGATGCTCGAG 129129

QY 841 AGCTTTAAGCGCTGACTCACCGCGAGGGGAAATGGAAGGGCTAAGGAGCGCCTTC 900

DB 129128 AGCTTTAAGCGCTGACTCACCGCGAGGGGAAATGGAAGGGCTAAGGAGCGCCTTC 129069

QY 901 GAAGCTGCTTTCATCAATGATTTAGAGCAAGTGTCCAGTTGATGGGTTTCATGAAGCT 960

DB 129068 GAAGCTGCTTTCATCAATGATTTAGAGCAAGTGTCCAGTTGATGGGTTTCATGAAGCT 129009

QY 961 ATATTAAACCATGTTAGAACCAATCACTTTTACTTAAGTACTTCCATAGGTCACGATGT 1020

DB 129008 ATATTAAACCATGTTAGAACCAATCACTTTTACTTAAGTACTTCCATAGGTCACGATGT 128949

QY 1021 GATCATGGAATCTTTCATACAGGTCCTCTTTTGGGGGCGAGCTCTTTTACTGTCCATCGG 1080

DB 128948 GATCATGGAATCTTTCATACAGGTCCTCTTTTGGGGGCGAGCTCTTTTACTGTCCATCGG 128889

QY 1081 ACCGCAAGATGACTGTGTGATTAACAGAGAAATTAAGCGCGAAGGACTCATTTGCGGTCT 1140

DB 128888 ACCGCAAGATGACTGTGTGATTAACAGAGAAATTAAGCGCGAAGGACTCATTTGCGGTCT 128829

QY 1141 TCTCGTGTGTTAAATTTCTGACGCTCTTTTGTTCATCGCGGCACTTTGCGCGTGTGATCT 1200

DB 128828 TCTCGTGTGTTAAATTTCTGACGCTCTTTTGTTCATCGCGGCACTTTGCGCGTGTGATCT 128769

QY 1201 TTTCTCAATCGCGCGCGATCGTGTTCGATTAATATATGCGTGGGTTGGCATCGCTTACCT 1260

DB 128768 TTTCTCAATCGCGCGCGATCGTGTTCGATTAATATATGCGTGGGTTGGCATCGCTTACCT 128709

QY 1261 GTTATGTTTGGCGTCAATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1320

DB 128708 GTTATGTTTGGCGTCAATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 128649

QY 1321 GATCATTTAGAAACAGAAACCAACCGTCCCATGACACGCTTTTGGCGGTTCGCGGT 1380

DB 128648 GATCATTTAGAAACAGAAACCAACCGTCCCATGACACGCTTTTGGCGGTTCGCGGT 128589

QY 1381 GGCCACTGACACGCGCAACCGGCTGGGTTGAGGTGAGCGTTCGATTAAGCAGCGGTTTG 1440

DB 128588 GGCCACTGACACGCGCAACCGGCTGGGTTGAGGTGAGCGTTCGATTAAGCAGCGGTTTG 128529

QY 1441 GGTAAAGCCCATGTTGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500

DB 128528 GGTAAAGCCCATGTTGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 128469

QY 1501 CGCGTTCGCTGTTTATCGCGCGCTCGCGCGCAATACGCGGCAACCGGACGCTGCTGCTGCT 1560

DB 128468 CGCGTTCGCTGTTTATCGCGCGCTCGCGCGCAATACGCGGCAACCGGACGCTGCTGCTGCT 128409

QY 1561 CGCGCTGGCGCTTCGCGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620

DB 128408 CGCGCTGGCGCTTCGCGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 128349

QY 1621 AGCATTTGACGCGCGCTTCGCGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680

DB 128348 AGCATTTGACGCGCGCTTCGCGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 128289

QY 1681 AGTTGTGATGACCGCATTCGCGCATCAAACTCATGTTGATGGGTTAGTTTTCGCGGCTTT 1740

DB 128288 AGTTGTGATGACCGCATTCGCGCATCAAACTCATGTTGATGGGTTAGTTTTCGCGGCTTT 128229

QY 1741 GGAATCGGTGGCTTCGCGCAAAATGTTGATCGCGGCTGCTGCGAAATCTCATCGATCGC 1800

DB 128228 GGAATCGGTGGCTTCGCGCAAAATGTTGATCGCGGCTGCTGCGAAATCTCATCGATCGC 128169

QY 1801 CTCCAACTCGCGCTCAGAAATCCCAAGTGTGATGATCAAGGCTGTTGTTCCAGCTG 1860

DB 128168 CTCCAACTCGCGCTCAGAAATCCCAAGTGTGATGATCAAGGCTGTTGTTCCAGCTG 128109

QY 1861 CTCCAACTCGAGCAACCAATCAATGCACTGCTGACGCTGCTGCGGCTGCTGCTGCTGCTG 1920



Thu Mar 18 12:30:59 2004

128108 CTCACTGACGAGCACCACCAATCAATGACACTGCTCAAGTATCCCGCGCGTACTCTCTCTTG 128049  
 1921 CTCGCGCAGACCCATGCTAGCGGCATCTGCGAAGTGAATGCTGCGCGGTTCTCTGGGCGAT 1980  
 128048 CTCGCGCAGACCCATGCTAGCGGCATCTGCGAAGTGAATGCTGCGCGGTTCTCTGGGCGAT 127989  
 1981 GTCAATGAGCTGCGGACCAATATCAATATTTTCAAGTCAATGCTGCGGTTCTCTGGGCGAT 2040  
 127988 GTCAATGAGCTGCGGACCAATATCAATATTTTCAAGTCAATGCTGCGGTTCTCTGGGCGAT 127929  
 2041 CTTACCTGCTGCGGCGGAGACCTCTGGAATTCATGAGATATTTGTCGCGGAGCAG 2100  
 127928 CTTACCTGCTGCGGCGGAGACCTCTGGAATTCATGAGATATTTGTCGCGGAGCAG 127869  
 2101 GCCCTGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCAATTTTGGCAGCTGACTGGA 2160  
 127868 GCCCTGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCAATTTTGGCAGCTGACTGGA 127809  
 2161 CAAGTCTTCAACGCTCATCGCCGCTTCTCCACCCAGCAATTAATGATGAATGCTGAG 2220  
 127808 CAAGTCTTCAACGCTCATCGCCGCTTCTCCACCCAGCAATTAATGATGAATGCTGAG 127749  
 2221 CTGATGAATCAGAGCGGCGAGCCCTCCCGCATGAATCTCAGCGCGCTCCGCTGTGAG 2280  
 127748 CTGATGAATCAGAGCGGCGAGCCCTCCCGCATGAATCTCAGCGCGCTCCGCTGTGAG 127689  
 2281 CTTCTGACCGTAGGAGAAATACCCAGTAAAGAGCCTTCCAGAGCGCAATGCTCAGC 2340  
 127688 CTTCTGACCGTAGGAGAAATACCCAGTAAAGAGCCTTCCAGAGCGCAATGCTCAGC 127629  
 2341 CAATCGCTACATGCTTTCTTCTTCAAGAGGATCT 2374  
 127628 CAATCGCTACATGCTTTCTTCTTCAAGAGGATCT 127595

RESULT 4  
 ID AAF71777/c  
 XX AAF71777 standard; DNA; 993 BP.  
 AC AAF71777;  
 XX  
 DT 30-APR-2001 (first entry)  
 XX  
 XX Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:49.  
 DE  
 XX Corynebacterium glutamicum; metabolic pathway protein; MP protein;  
 KW fine chemical production; microorganism; organic acid; nucleoside;  
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;  
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
 KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN W0200100843-A2.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 23-JUN-2000; 2000WO-IB000923.  
 XX  
 XX 25-JUN-1999; 99US-0141031P.  
 PR 01-JUL-1999; 99DE-01030476.  
 PR 02-JUL-1999; 99US-0142101P.  
 PR 08-JUL-1999; 99DE-01031415.  
 PR 08-JUL-1999; 99DE-01031418.  
 PR 08-JUL-1999; 99DE-01031419.  
 PR 08-JUL-1999; 99DE-01031420.  
 PR 08-JUL-1999; 99DE-01031424.  
 PR 08-JUL-1999; 99DE-01031428.  
 PR 08-JUL-1999; 99DE-01031434.  
 PR 08-JUL-1999; 99DE-01031435.  
 PR 08-JUL-1999; 99DE-01031443.  
 PR 08-JUL-1999; 99DE-01031453.  
 PR 08-JUL-1999; 99DE-01031457.

08-JUL-1999; 99DE-01031465.  
 08-JUL-1999; 99DE-01031478.  
 08-JUL-1999; 99DE-01031510.  
 08-JUL-1999; 99DE-01031541.  
 08-JUL-1999; 99DE-01031573.  
 08-JUL-1999; 99DE-01031592.  
 08-JUL-1999; 99DE-01031632.  
 08-JUL-1999; 99DE-01031634.  
 08-JUL-1999; 99DE-01031636.  
 08-JUL-1999; 99DE-01031638.  
 08-JUL-1999; 99DE-01031640.  
 08-JUL-1999; 99DE-01031642.  
 08-JUL-1999; 99DE-01031644.  
 08-JUL-1999; 99DE-01031646.  
 08-JUL-1999; 99DE-01031648.  
 08-JUL-1999; 99DE-01031650.  
 08-JUL-1999; 99DE-01031652.  
 08-JUL-1999; 99DE-01031654.  
 08-JUL-1999; 99DE-01031656.  
 08-JUL-1999; 99DE-01031658.  
 08-JUL-1999; 99DE-01031660.  
 08-JUL-1999; 99DE-01031662.  
 08-JUL-1999; 99DE-01031664.  
 08-JUL-1999; 99DE-01031666.  
 08-JUL-1999; 99DE-01031668.  
 08-JUL-1999; 99DE-01031670.  
 08-JUL-1999; 99DE-01031672.  
 08-JUL-1999; 99DE-01031674.  
 08-JUL-1999; 99DE-01031676.  
 08-JUL-1999; 99DE-01031678.  
 08-JUL-1999; 99DE-01031680.  
 08-JUL-1999; 99DE-01031682.  
 08-JUL-1999; 99DE-01031684.  
 08-JUL-1999; 99DE-01031686.  
 08-JUL-1999; 99DE-01031688.  
 08-JUL-1999; 99DE-01031690.  
 08-JUL-1999; 99DE-01031692.  
 08-JUL-1999; 99DE-01031694.  
 08-JUL-1999; 99DE-01031696.  
 08-JUL-1999; 99DE-01031698.  
 08-JUL-1999; 99DE-01031700.  
 08-JUL-1999; 99DE-01031702.  
 08-JUL-1999; 99DE-01031704.  
 08-JUL-1999; 99DE-01031706.  
 08-JUL-1999; 99DE-01031708.  
 08-JUL-1999; 99DE-01031710.  
 08-JUL-1999; 99DE-01031712.  
 08-JUL-1999; 99DE-01031714.  
 08-JUL-1999; 99DE-01031716.  
 08-JUL-1999; 99DE-01031718.  
 08-JUL-1999; 99DE-01031720.  
 08-JUL-1999; 99DE-01031722.  
 08-JUL-1999; 99DE-01031724.  
 08-JUL-1999; 99DE-01031726.  
 08-JUL-1999; 99DE-01031728.  
 08-JUL-1999; 99DE-01031730.  
 08-JUL-1999; 99DE-01031732.  
 08-JUL-1999; 99DE-01031734.  
 08-JUL-1999; 99DE-01031736.  
 08-JUL-1999; 99DE-01031738.  
 08-JUL-1999; 99DE-01031740.  
 08-JUL-1999; 99DE-01031742.  
 08-JUL-1999; 99DE-01031744.  
 08-JUL-1999; 99DE-01031746.  
 08-JUL-1999; 99DE-01031748.  
 08-JUL-1999; 99DE-01031750.  
 08-JUL-1999; 99DE-01031752.  
 08-JUL-1999; 99DE-01031754.  
 08-JUL-1999; 99DE-01031756.  
 08-JUL-1999; 99DE-01031758.  
 08-JUL-1999; 99DE-01031760.  
 08-JUL-1999; 99DE-01031762.  
 08-JUL-1999; 99DE-01031764.  
 08-JUL-1999; 99DE-01031766.  
 08-JUL-1999; 99DE-01031768.  
 08-JUL-1999; 99DE-01031770.  
 08-JUL-1999; 99DE-01031772.  
 08-JUL-1999; 99DE-01031774.  
 08-JUL-1999; 99DE-01031776.  
 08-JUL-1999; 99DE-01031778.  
 08-JUL-1999; 99DE-01031780.  
 08-JUL-1999; 99DE-01031782.  
 08-JUL-1999; 99DE-01031784.  
 08-JUL-1999; 99DE-01031786.  
 08-JUL-1999; 99DE-01031788.  
 08-JUL-1999; 99DE-01031790.  
 08-JUL-1999; 99DE-01031792.  
 08-JUL-1999; 99DE-01031794.  
 08-JUL-1999; 99DE-01031796.  
 08-JUL-1999; 99DE-01031798.  
 08-JUL-1999; 99DE-01031800.  
 08-JUL-1999; 99DE-01031802.  
 08-JUL-1999; 99DE-01031804.  
 08-JUL-1999; 99DE-01031806.  
 08-JUL-1999; 99DE-01031808.  
 08-JUL-1999; 99DE-01031810.  
 08-JUL-1999; 99DE-01031812.  
 08-JUL-1999; 99DE-01031814.  
 08-JUL-1999; 99DE-01031816.  
 08-JUL-1999; 99DE-01031818.  
 08-JUL-1999; 99DE-01031820.  
 08-JUL-1999; 99DE-01031822.  
 08-JUL-1999; 99DE-01031824.  
 08-JUL-1999; 99DE-01031826.  
 08-JUL-1999; 99DE-01031828.  
 08-JUL-1999; 99DE-01031830.  
 08-JUL-1999; 99DE-01031832.  
 08-JUL-1999; 99DE-01031834.  
 08-JUL-1999; 99DE-01031836.  
 08-JUL-1999; 99DE-01031838.  
 08-JUL-1999; 99DE-01031840.  
 08-JUL-1999; 99DE-01031842.  
 08-JUL-1999; 99DE-01031844.  
 08-JUL-1999; 99DE-01031846.  
 08-JUL-1999; 99DE-01031848.  
 08-JUL-1999; 99DE-01031850.  
 08-JUL-1999; 99DE-01031852.  
 08-JUL-1999; 99DE-01031854.  
 08-JUL-1999; 99DE-01031856.  
 08-JUL-1999; 99DE-01031858.  
 08-JUL-1999; 99DE-01031860.  
 08-JUL-1999; 99DE-01031862.  
 08-JUL-1999; 99DE-01031864.  
 08-JUL-1999; 99DE-01031866.  
 08-JUL-1999; 99DE-01031868.  
 08-JUL-1999; 99DE-01031870.  
 08-JUL-1999; 99DE-01031872.  
 08-JUL-1999; 99DE-01031874.  
 08-JUL-1999; 99DE-01031876.  
 08-JUL-1999; 99DE-01031878.  
 08-JUL-1999; 99DE-01031880.  
 08-JUL-1999; 99DE-01031882.  
 08-JUL-1999; 99DE-01031884.  
 08-JUL-1999; 99DE-01031886.  
 08-JUL-1999; 99DE-01031888.  
 08-JUL-1999; 99DE-01031890.  
 08-JUL-1999; 99DE-01031892.  
 08-JUL-1999; 99DE-01031894.  
 08-JUL-1999; 99DE-01031896.  
 08-JUL-1999; 99DE-01031898.  
 08-JUL-1999; 99DE-01031900.  
 08-JUL-1999; 99DE-01031902.  
 08-JUL-1999; 99DE-01031904.  
 08-JUL-1999; 99DE-01031906.  
 08-JUL-1999; 99DE-01031908.  
 08-JUL-1999; 99DE-01031910.  
 08-JUL-1999; 99DE-01031912.  
 08-JUL-1999; 99DE-01031914.  
 08-JUL-1999; 99DE-01031916.  
 08-JUL-1999; 99DE-01031918.  
 08-JUL-1999; 99DE-01031920.  
 08-JUL-1999; 99DE-01031922.  
 08-JUL-1999; 99DE-01031924.  
 08-JUL-1999; 99DE-01031926.  
 08-JUL-1999; 99DE-01031928.  
 08-JUL-1999; 99DE-01031930.  
 08-JUL-1999; 99DE-01031932.  
 08-JUL-1999; 99DE-01031934.  
 08-JUL-1999; 99DE-01031936.  
 08-JUL-1999; 99DE-01031938.  
 08-JUL-1999; 99DE-01031940.  
 08-JUL-1999; 99DE-01031942.  
 08-JUL-1999; 99DE-01031944.  
 08-JUL-1999; 99DE-01031946.  
 08-JUL-1999; 99DE-01031948.  
 08-JUL-1999; 99DE-01031950.  
 08-JUL-1999; 99DE-01031952.  
 08-JUL-1999; 99DE-01031954.  
 08-JUL-1999; 99DE-01031956.  
 08-JUL-1999; 99DE-01031958.  
 08-JUL-1999; 99DE-01031960.  
 08-JUL-1999; 99DE-01031962.  
 08-JUL-1999; 99DE-01031964.  
 08-JUL-1999; 99DE-01031966.  
 08-JUL-1999; 99DE-01031968.  
 08-JUL-1999; 99DE-01031970.  
 08-JUL-1999; 99DE-01031972.  
 08-JUL-1999; 99DE-01031974.  
 08-JUL-1999; 99DE-01031976.  
 08-JUL-1999; 99DE-01031978.  
 08-JUL-1999; 99DE-01031980.  
 08-JUL-1999; 99DE-01031982.  
 08-JUL-1999; 99DE-01031984.  
 08-JUL-1999; 99DE-01031986.  
 08-JUL-1999; 99DE-01031988.  
 08-JUL-1999; 99DE-01031990.  
 08-JUL-1999; 99DE-01031992.  
 08-JUL-1999; 99DE-01031994.  
 08-JUL-1999; 99DE-01031996.  
 08-JUL-1999; 99DE-01031998.  
 08-JUL-1999; 99DE-01032000.

(BADI ) BASF AG.

Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;

WPI; 2001-137957/14.

P-PSDB; AAB79658.

Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway proteins, useful for producing fine chemicals in microorganisms, including organic acids, nonproteinogenic amino acids, and purine and pyrimidine bases.

Claim 3; Page 226-228; 1737pp; English.

AAE71753 to AAE72330 encode the Corynebacterium glutamicum metabolic pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP nucleic acids are useful for the production of fine chemicals in microorganisms, including organic acids, nonproteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, polyketides and enzymes

Sequence 993 BP; 233 A; 256 C; 268 G; 236 T; 0 U; 0 Other;

Query Match 41.8%; Score 993; DB 4; Length 993;  
 Best Local Similarity 100.0%; Pred.No. 4.9e-299;  
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 62 CTGAACCTTTTTCAGAGTAACCTAAGCGCGCAATCCCTCGATTCGTCATCAACGCGCG 121  
 Db 993 CTGAACCTTTTTCAGAGTAACCTAAGCGCGCAATCCCTCGATTCGTCATCAACGCGCG 934  
 QY 122 TCTGTGAGTCTAGCTAGAGTCTAGATCCAGCGCCCATCTGTCGCAATACATCGGTGTG 181  
 Db 933 TCTGTGAGTCTAGCTAGAGTCTAGATCCAGCGCCCATCTGTCGCAATACATCGGTGTG 874  
 QY 182 TCAATGGGTATCTCATCGAGGAGGATCACTTCTCTGCTTTTGTAGCATGGAGGAGCTTGG 241  
 Db 873 TCAATGGGTATCTCATCGAGGAGGATCACTTCTCTGCTTTTGTAGCATGGAGGAGCTTGG 814  
 QY 242 GTTTCGGGAAGAGTCCCAACCAAGCGCTCGGCGAATTCCTCAGCAAACTTCGCC 301  
 Db 813 GTTTCGGGAAGAGTCCCAACCAAGCGCTCGGCGAATTCCTCAGCAAACTTCGCC 754  
 QY 302 GACGGGACAATGGATAGCGCTCGGCCACACAGGACCATCGAGCGCCGCTCCAGGTCA 361  
 Db 753 GACGGGACAATGGATAGCGCTCGGCCACACAGGACCATCGAGCGCCGCTCCAGGTCA 694  
 QY 362 CGGTCTTGAACACATCTTTGGGACCGCAAGCGTAAAGCGGCAATCGAGCGCCCAATCTAGT 421  
 Db 693 CGGTCTTGAACACATCTTTGGGACCGCAAGCGTAAAGCGGCAATCGAGCGCCCAATCTAGT 634  
 QY 422 TTCCCATCAACCATGTAGGCATCCCGCAATGAGGGGGTTGCAATGGCCAGTGGCGCATG 481  
 Db 633 TTCCCATCAACCATGTAGGCATCCCGCAATGAGGGGGTTGCAATGGCCAGTGGCGCATG 574  
 QY 482 GTTCCAGTCTTACTTCACTTCACTCCCGCCACGCGATAGCTTCACTCGGCTTACCGCTCT 541  
 Db 573 GTTCCAGTCTTACTTCACTTCACTCCCGCCACGCGATAGCTTCACTCGGCTTACCGCTCT 514  
 QY 542 AAAACATCTCCAGCGCGCAGCAAGATATGTGCGCTTCACTTCCAGCGCGAGCGTG 601  
 Db 513 AAAACATCTCCAGCGCGCAGCAAGATATGTGCGCTTCACTTCCAGCGCGAGCGTG 454  
 QY 602 AGCTGTGCTCCACCCCAAGAGTACCTGTTGAACACGCGGAGGAAACCATGTGATAGC 661  
 Db 453 AGCTGTGCTCCACCCCAAGAGTACCTGTTGAACACGCGGAGGAAACCATGTGATAGC 394  
 QY 662 GAATCTGCTGTATGATGCGATGTTAAACGGGATTTCAAGCGGCTCCAGATAGTTCGCT 721  
 Db 393 GAATCTGCTGTATGATGCGATGTTAAACGGGATTTCAAGCGGCTCCAGATAGTTCGCT 334  
 QY 722 TTAGTTTCTGCTTGAACCAACCACTTTCCGCGCTGTTGCAACAGGACTTCACCGCT 781  
 Db 333 TTAGTTTCTGCTTGAACCAACCACTTTCCGCGCTGTTGCAACAGGACTTCACCGCT 274  
 QY 782 TCGGTGTTGTTGCGCGGTGGTGCGATACCAACCTCCGACCCAGCTGATGCTCGAGA 841  
 Db 273 TCGGTGTTGTTGCGCGGTGGTGCGATACCAACCTCCGACCCAGCTGATGCTCGAGA 214  
 QY 842 GCTTTAAGCGCTGATCTACCCGCGAGGGGAAATCGAAAGGGCTTAAGGAGCGCTTCG 901  
 Db 213 GCTTTAAGCGCTGATCTACCCGCGAGGGGAAATCGAAAGGGCTTAAGGAGCGCTTCG 154  
 QY 902 AAGCTGCTCTCATCAATGATTTGAGAGCAAGGTGTCAGTTGAATGGGGTTTCATGAGCTA 961  
 Db 153 AAGCTGCTCTCATCAATGATTTGAGAGCAAGGTGTCAGTTGAATGGGGTTTCATGAGCTA 94  
 QY 962 TATTAAACCATGTTAAGACCAATCATTTTACTTACTTACTTCCATAGCTCAGATGGTG 1021  
 Db 93 TATTAAACCATGTTAAGACCAATCATTTTACTTACTTACTTCCATAGCTCAGATGGTG 34  
 QY 1022 ATCATGGAATCTTCAATACAGCTCTGCTTTTG 1054  
 Db 33 ATCATGGAATCTTCAATACAGCTCTGCTTTTG 1

RESULT 5  
 AAS96096/c  
 ID AAS96096 standard; DNA; 993 BP.  
 XX

AC AAS96096;  
 XX 26-FEB-2002 (first entry)  
 DT C. glutamicum gene #21 encoding metabolic pathway protein.  
 XX  
 DE Metabolic pathway protein; MP; lysine biosynthesis pathway;  
 XX methionine biosynthesis pathway; large-scale production of fine chemical;  
 KW Corynebacterium diphtheriae; diphtheria; ds.  
 KW  
 XX Corynebacterium glutamicum.  
 OS  
 XX W0200166573-A2.  
 PN 13-SEP-2001.  
 XX 22-DEC-2000; 2000WO-IB002035.  
 PF 09-MAR-2000; 2000US-0187970P.  
 PR 23-JUN-2000; 2000US-00606740.  
 XX (BADI) BASF AG.  
 PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
 XX Kim J, Lee H, Hwang B;  
 PI WPI; 2001-582269/65.  
 DR P-PSDB; AAU71886.  
 XX Nucleic acids encoding metabolic pathway proteins from Corynebacterium  
 PT glutamicum, useful for producing methionine and lysine in Corynebacterium  
 PT and Brevibacterium.  
 XX Disclosure; Page 207-208; 316pp; English.  
 PS The present invention relates to the isolation of novel Corynebacterium  
 XX glutamicum genes encoding metabolic pathway (MP) proteins (AAU71863-  
 CC AAU71922). The metabolic pathway proteins of the invention include  
 CC enzymes involved in the lysine and methionine biosynthetic pathways. The  
 CC polynucleotide sequences of the invention can be used for the large-scale  
 CC production and/or modulation of expression of fine chemicals such as  
 CC lysine and methionine. The sequences of the invention may be used to  
 CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a  
 CC subject to detect diphtheria. AAS96073-AAS96132 represent C. glutamicum  
 CC genes encoding the novel metabolic pathway proteins of the invention  
 XX  
 XX Sequence 993 BP; 233 A; 256 C; 268 G; 236 T; 0 U; 0 Other;  
 SQ  
 Query Match 41.8%; Score 993; DB 4; Length 993;  
 Best Local Similarity 100.0%; Pred. No. 4,9e-299;  
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 62 CTGAACCTTTTTCAGAGTAACCTAAGCGCGCAATCCCTCGATTCGTCATCAACGCGCG 121  
 Db 993 CTGAACCTTTTTCAGAGTAACCTAAGCGCGCAATCCCTCGATTCGTCATCAACGCGCG 934  
 QY 122 TCTGTGAGTCTAGCTAGAGTCTAGATCCAGCGCCCATCTGTCGCAATACATCGGTGTG 181  
 Db 933 TCTGTGAGTCTAGCTAGAGTCTAGATCCAGCGCCCATCTGTCGCAATACATCGGTGTG 874  
 QY 182 TCAATGGGTATCTCATCGAGGAGGATCACTTCTCTGCTTTTGTAGCATGGAGGAGCTTGG 241  
 Db 873 TCAATGGGTATCTCATCGAGGAGGATCACTTCTCTGCTTTTGTAGCATGGAGGAGCTTGG 814  
 QY 242 GTTTCGGGAAGAGTCCCAACCAAGCGCTCGGCGAATTCCTCAGCAAACTTCGCC 301  
 Db 813 GTTTCGGGAAGAGTCCCAACCAAGCGCTCGGCGAATTCCTCAGCAAACTTCGCC 754  
 QY 302 GACGGGACAATGGATAGCGCTCGGCCACACAGGACCATCGAGCGCCGCTCCAGGTCA 361  
 Db 753 GACGGGACAATGGATAGCGCTCGGCCACACAGGACCATCGAGCGCCGCTCCAGGTCA 694  
 QY 362 CGGTCTTGAACACATCTTTGGGACCGCAAGCGTAAAGCGGCAATCGAGCGCCCAATCTAGT 421

RESULT 6	
AAAH68421/c	
ID AAH68421	standard; DNA; 870 BP.
XX	
AC	AAH68421;
XX	
DT	26-SEP-2001 (first entry)
XX	
DE	C glutamicum coding sequence fragment SEQ ID NO: 3456.
DE	
XX	Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW	organic acid synthesis; ds.
KW	
XX	
OS	Corynebacterium glutamicum.
XX	
PN	EP1108790-A2.
XX	
PD	20-JUN-2001.
XX	
PF	18-DEC-2000; 2000EP-00127688.
XX	
XX	16-DEC-1999; 99JP-00377484.
PR	07-APR-2000; 2000JP-00159162.
PR	03-AUG-2000; 2000FP-00280988.
XX	
XX	

(KYOW) KIOWA HAKKO KOGYO KK.

Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
Tateishi N, Senoh A, Ikeda M, Ozaki A;  
WPI: 2001-376931/40.  
P-PSDB: AAG3202.

Novel polynucleotides derived from Coryneform bacteria, for identifying mutation point of a gene, measuring expression of a gene, analyzing expression profile or pattern of a gene and identifying homologous gene.

Claim 1; SEQ ID NO 3456; 246pp + Sequence Listing; English.

The present invention provides a number of nucleotide and protein sequences from the Coryneform bacterium *Corynebacterium glutamicum*. These are useful for identifying the mutation point of a gene derived from a mutant of coryneform bacterium, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from coryneform bacterium, and identifying a homologue of a gene derived from coryneform bacterium. Coryneform bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a nucleic acid described in the exemplification of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office

Sequence 870 BP; 192 A; 236 C; 245 G; 197 T; 0 U; 0 Other;

Query Match 36.6%; Score 870; DB 5; Length 870;  
Best Local Similarity 100.0%; Pred. No. 1.3e-260; Indels 0; Gaps 0;  
Matches 870; Conservative 0; Mismatches 0;

85 AGCGCGCAATCCCTCGATTGCTGATCAACGACGGCGTCTGTGAGTCTAGCTAGAGATCT 144  
870 AGCGCGCAATCCCTCGATTGCTGATCAACGACGGCGTCTGTGAGTCTAGCTAGAGATCT 811

145 AGATTTCAGGCGGCATCCGTTGCCAATACATCGGTGTGTAATGGGTATCTCATCGAGGAG 204  
810 AGATTTCAGGCGGCATCCGTTGCCAATACATCGGTGTGTAATGGGTATCTCATCGAGGAG 751

205 GATCACTTCTCTGTCTTTAGCATGGGACGAGCTTGGGTTTCGGGAAGAGTCCCAACC 264  
750 GATCACTTCTCTGTCTTTAGCATGGGACGAGCTTGGGTTTCGGGAAGAGTCCCAACC 691

265 AAGGCTTCGGGAATGCTCTACCAAACTTCGCGCGACGGGAACAATGGATACGGCCT 324  
690 AAGGCTTCGGGAATGCTCTACCAAACTTCGCGCGACGGGAACAATGGATACGGCCT 631

325 GCGCCCCACAGGACCAATCGACGCGCCGTCACAGTCAAGTCTTGAACCACTTTGGG 384  
630 GCGCCCCACAGGACCAATCGACGCGCCGTCACAGTCAAGTCTTGAACCACTTTGGG 571

385 ACCGAAGCGTAAGACGGGCAATCGCAGCCCAATCTAGTTTCCCATCAACCATGTAGGCA 444  
570 ACCGAAGCGTAAGACGGGCAATCGCAGCCCGTCCAGTCAAGTCTTGAACCACTTTGGG 511

445 CCGCAATGAGGGGTTCGAATGGCCAGTGGCGCATGGTTCGAAGTTCTACTCTTCA 504  
510 CCGCAATGAGGGGTTCGAATGGCCAGTGGCGCATGGTTCGAAGTTCTACTCTTCA 451

505 TCCGCCACGCGAATAGCTTTCACGGGTTACCGTCTCTAAACATCTCCACGCGCGACGA 564  
450 TCCGCCACGCGAATAGCTTTCACGGGTTACCGTCTCTAAACATCTCCACGCGCGACGA 391

565 GGAATAATGTGTGGCTTCATCTTCCAAAGCGGAGTGTAGCGTCTCTCCACCCCAAGAGC 624  
390 GGAATAATGTGTGGCTTCATCTTCCAAAGCGGAGTGTAGCGTCTCTCCACCCCAAGAGC 331

625 TACTCTGTTGACACGGGAGAAAACCATGTGGATAGCGAATCTCGTGTGATGGCGATGGT 684  
330 TACTCTGTTGACACGGGAGAAAACCATGTGGATAGCGAATCTCGTGTGATGGCGATGGT 271

QY 685 TAACGGGATTCAGCAGCGCTCCAGATAGTTGGCTTTAGTTTCTGCTGCAGCACAC 744  
 DB 270 TAACGGGATTCAGCAGCGCTCCAGATAGTTGGCTTTAGTTTCTGCTGCAGCACAC 211  
 QY 745 CATTTTCCGGCTGCTTCACAGGACTTCACCGCTTCGGTTGCTTGGCCGGTGGGT 804  
 DB 210 CATTTTCCGGCTGCTTCACAGGACTTCACCGCTTCGGTTGCTTGGCCGGTGGGT 151  
 QY 805 GCGGATACCAACACTTCGACCCACGTCGATGCTCGAGAGCTTTAACGGCTGACTCAGCGC 864  
 DB 150 GCGGATACCAACACTTCGACCCACGTCGATGCTCGAGAGCTTTAACGGCTGACTCAGCGC 91  
 QY 865 CGAGGGGAATCGAAGGCTTAAGAGCGCTTCGAGCTGCTTCATCAATGATTGA 924  
 DB 90 CGAGGGGAATCGAAGGCTTAAGAGCGCTTCGAGCTGCTTCATCAATGATTGA 31  
 QY 925 GAGCAAGTGTCCAGTTGAATGGGTTTCAT 954  
 DB 30 GAGCAAGTGTCCAGTTGAATGGGTTTCAT 1

RESULT 7  
 ID AAF71779 standard; DNA; 822 BP.  
 AC AAF71779;  
 DT 30-APR-2001 (first entry)  
 XX Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:53.  
 DE  
 XX  
 KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;  
 KW fine chemical production; microorganism; organic acid; nucleoside;  
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;  
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
 KW carbonyl; aromatic compound; cofactor; polyketide; enzyme; ds.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN WO200100843-A2.  
 XX  
 XX 04-JAN-2001.  
 XX  
 XX 23-JUN-2000; 2000WO-IB000923.  
 XX  
 PR 25-JUN-1999; 99US-0141031P.  
 PR 01-JUL-1999; 99DE-01030476.  
 PR 02-JUL-1999; 99US-0142101P.  
 PR 08-JUL-1999; 99DE-01031415.  
 PR 08-JUL-1999; 99DE-01031418.  
 PR 08-JUL-1999; 99DE-01031419.  
 PR 08-JUL-1999; 99DE-01031420.  
 PR 08-JUL-1999; 99DE-01031424.  
 PR 08-JUL-1999; 99DE-01031428.  
 PR 08-JUL-1999; 99DE-01031434.  
 PR 08-JUL-1999; 99DE-01031435.  
 PR 08-JUL-1999; 99DE-01031443.  
 PR 08-JUL-1999; 99DE-01031453.  
 PR 08-JUL-1999; 99DE-01031457.  
 PR 08-JUL-1999; 99DE-01031465.  
 PR 08-JUL-1999; 99DE-01031478.  
 PR 08-JUL-1999; 99DE-01031510.  
 PR 08-JUL-1999; 99DE-01031541.  
 PR 08-JUL-1999; 99DE-01031573.  
 PR 08-JUL-1999; 99DE-01031592.  
 PR 08-JUL-1999; 99DE-01031632.  
 PR 08-JUL-1999; 99DE-01031634.  
 PR 08-JUL-1999; 99DE-01031636.  
 PR 09-JUL-1999; 99DE-01032125.  
 PR 09-JUL-1999; 99DE-01032126.  
 PR 09-JUL-1999; 99DE-01032130.  
 PR 09-JUL-1999; 99DE-01032186.  
 PR 09-JUL-1999; 99DE-01032206.

PR 09-JUL-1999; 99DE-01032227.  
 PR 09-JUL-1999; 99DE-01032228.  
 PR 09-JUL-1999; 99DE-01032229.  
 PR 09-JUL-1999; 99DE-01032230.  
 PR 14-JUL-1999; 99DE-01032922.  
 PR 14-JUL-1999; 99DE-01032926.  
 PR 14-JUL-1999; 99DE-01032928.  
 PR 14-JUL-1999; 99DE-01033004.  
 PR 14-JUL-1999; 99DE-01033005.  
 PR 12-AUG-1999; 99DE-01033006.  
 PR 12-AUG-1999; 99US-0148613P.  
 PR 27-AUG-1999; 99DE-01040764.  
 PR 27-AUG-1999; 99DE-01040765.  
 PR 27-AUG-1999; 99DE-01040766.  
 PR 27-AUG-1999; 99DE-01040832.  
 PR 31-AUG-1999; 99DE-01041378.  
 PR 31-AUG-1999; 99DE-01041379.  
 PR 31-AUG-1999; 99DE-01041380.  
 PR 31-AUG-1999; 99DE-01041394.  
 PR 31-AUG-1999; 99DE-01041396.  
 PR 03-SEP-1999; 99DE-01042076.  
 PR 03-SEP-1999; 99DE-01042077.  
 PR 03-SEP-1999; 99DE-01042079.  
 PR 03-SEP-1999; 99DE-01042086.  
 PR 03-SEP-1999; 99DE-01042087.  
 PR 03-SEP-1999; 99DE-01042088.  
 PR 03-SEP-1999; 99DE-01042095.  
 PR 03-SEP-1999; 99DE-01042124.  
 PR 03-SEP-1999; 99DE-01042129.  
 PR 09-MAR-2000; 2000US-0187970P.  
 XX  
 PA (BADI ) BASF AG.  
 XX  
 PI Pompejus M, Kroege B, Schroeder H, Zelder O, Haberhauer G;  
 XX WPI; 2001-137957/14.  
 DR P-PSDB; AAB79660.  
 XX  
 PT Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway  
 PT proteins, useful for producing fine chemicals in microorganisms,  
 PT including organic acids, nonproteinogenic amino acids, and purine and  
 PT pyrimidine bases.  
 XX  
 PS Claim 3; Page 233-234; 1737pp; English.  
 XX  
 CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic  
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP  
 CC nucleic acids are useful for the production of fine chemicals in  
 CC microorganisms, including organic acids, nonproteinogenic amino acids,  
 CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated  
 CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,  
 CC vitamins, cofactors, polyketides and enzymes  
 XX  
 SQ Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;  
 Query Match 34.6%; Score 822; DB 4; Length 822;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-245;  
 Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 925 GAGCAAGTGTCCAGTTGAATGGGTTTCATGAGCTTATATTAACCATGTTAAGAACCA 984  
 DB 1 GAGCAAGTGTCCAGTTGAATGGGTTTCATGAGCTTATATTAACCATGTTAAGAACCA 60  
 QY 985 TCATTTTTAACTTCCATAGTTCAGATGGTGCATGCAATCTTCATTACAGG 1044  
 DB 61 TCATTTTTAACTTCCATAGTTCAGATGGTGCATGCAATCTTCATTACAGG 120  
 QY 1045 TCTGCTTTGGGGCCAGCTCTTTTACTGCTCAGCCGACGAGATGCTGCTGATTAA 1104  
 DB 121 TCTGCTTTGGGGCCAGCTCTTTTACTGCTCAGCCGACGAGATGCTGCTGATTAA 180  
 QY 1105 ACAAGGATTAAGCGCGAAGGACTCAATTGGGTTTCTTCGTGTTTAAATTCAGCT 1164

Thu Mar 18 12:30:59 2004

Db 181 ACAGGAATTAAAGCGGAGGACTCATTGGGTTCTTCTCGTGTTAAATTTCTGAGCT 240  
QY 1165 CTTTGTGTCTATCCGCGGACCTTGGGGTTGATCTTTGTCCTCAATGCGCGCGATCGT 1224  
Db 241 CTTTGTGTCTATCCGCGGACCTTGGGGTTGATCTTTGTCCTCAATGCGCGCGATCGT 300  
QY 1225 GCTCGATATTATGCGTGGGTTGGCATCGCTTACCTGTTATGTTTGGCGTATGCGAGC 1284  
Db 301 GCTCGATATTATGCGTGGGTTGGCATCGCTTACCTGTTATGTTTGGCGTATGCGAGC 360  
QY 1285 GAAAGACGCCATGACAAACCAAGGTGAAGCGCCACAGATCATGTGAAGAACAGACCAAC 1344  
Db 361 GAAAGACGCCATGACAAACCAAGGTGAAGCGCCACAGATCATGTGAAGAACAGACCAAC 420  
QY 1345 CGTGGCCGATGACAGCGCTTTGGCGGTTTCGCGGTCGCTGACAGCGCGACCGGCT 1404  
Db 421 CGTGGCCGATGACAGCGCTTTGGCGGTTTCGCGGTCGCTGACAGCGCGACCGGCT 480  
QY 1405 GCGGCTGGAGTGAAGCGTTCGATGAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCGCAT 1464  
Db 481 GCGGCTGGAGTGAAGCGTTCGATGAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCGCAT 540  
QY 1465 GGTCTGACCTGTTGAACCCGATGCTTATGACGCGCTTTGCTTTATCGCGCGCT 1524  
Db 541 GGTCTGACCTGTTGAACCCGATGCTTATGACGCGCTTTGCTTTATCGCGCGCT 600  
QY 1525 GCGCGCGATACGCGGACCGGACCGGATGCTTTTCCCGCTGGCGCTTCGCGGCAAG 1584  
Db 601 GCGCGCGATACGCGGACCGGACCGGATGCTTTTCCCGCTGGCGCTTCGCGGCAAG 660  
QY 1585 CCGATCTGCTTCCCGCTGGGTTTCGCGGACGAGCATGTGACGCGCGCTGTCAG 1644  
Db 661 CCGATCTGCTTCCCGCTGGGTTTCGCGGACGAGCATGTGACGCGCGCTGTCAG 720  
QY 1645 CCGCAAGTGTGGCGCTGATCAAGCTGCTGCGGAGTTGTGATGACGCGCATTTGGCCAT 1704  
Db 721 CCGCAAGTGTGGCGCTGATCAAGCTGCTGCGGAGTTGTGATGACGCGCATTTGGCCAT 780  
QY 1705 CAACCTGATGTTGATGGTTAGTTTTCGCGGTTTGGAAATC 1746  
Db 781 CAACCTGATGTTGATGGTTAGTTTTCGCGGTTTGGAAATC 822

RESULT 8  
AAS96098 standard; DNA; 822 BP.  
AC AAS96098;  
DT 26-FEB-2002 (first entry)  
XX C. glutamicum gene #23 encoding metabolic pathway protein.  
DE Metabolic pathway protein; MP; lysine biosynthesis pathway;  
KW methionine biosynthesis pathway; large-scale production of fine chemical;  
KW Corynebacterium diphtheriae; diphtheria; ds.  
XX Corynebacterium glutamicum.  
XX WO200166573-A2.  
XX 13-SEP-2001.  
XX 22-DEC-2000; 2000WO-IB002035.  
XX 09-MAR-2000; 2000US-0189797P.  
XX 23-JUN-2000; 2000US-00606740.  
XX (BADI ) BASF AG.  
XX Pompejus M, Kroeger B, Schroeder H, Zeider O, Habershauer G;  
PI Kim J, Lee H, Hwang B;  
XX

DR WPI: 2001-582269/65.  
DR P-PSDB; AAU71888.  
XX Nucleic acids encoding metabolic pathway proteins from Corynebacterium  
PT glutamicum, useful for producing methionine and lysine in Corynebacterium  
PT and Brevibacterium.  
XX Disclosure; Page 214-215; 316pp; English.  
XX The present invention relates to the isolation of novel Corynebacterium  
CC glutamicum genes encoding metabolic pathway (MP) proteins (AAU71883-  
CC AAU71922). The metabolic pathway proteins of the invention include  
CC enzymes involved in the lysine and methionine biosynthetic pathways. The  
CC polynucleotide sequences of the invention can be used for the large-scale  
CC production and/or modulation of expression of fine chemicals such as  
CC lysine and methionine. The sequences of the invention may be used to  
CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a  
CC subject to detect diphtheria. AAS96073-AAS96132 represent C. glutamicum  
CC genes encoding the novel metabolic pathway proteins of the invention  
XX Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;  
SQ Query Match 34.6%; Score 822; DB 4; Length 822;  
Best Local Similarity 100.0%; Pred. No. 1.2e-245;  
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 925 GAGCAAGTGTCCAGTTGAATGGGGTTCATGAAGCTATATTAACCATGTTAAGAACCA 984  
Db 1 GAGCAAGTGTCCAGTTGAATGGGGTTCATGAAGCTATATTAACCATGTTAAGAACCA 60  
QY 985 TCATTTTACTTAACTTCCATAGGTCATGATGATGATGATGATGATGATGATGATGAT 1044  
Db 61 TCATTTTACTTAACTTCCATAGGTCATGATGATGATGATGATGATGATGATGATGAT 120  
QY 1045 TCTGCTTTTGGGCGGAGCTCTTTTACTGTCATGCGGCGGAGATGATGATGATGATGAT 1104  
Db 121 TCTGCTTTTGGGCGGAGCTCTTTTACTGTCATGCGGCGGAGATGATGATGATGATGAT 180  
QY 1105 ACAAGAAATTAAGCGCGAGGACTCATTTGCGGTTCTTCTGCTGTTTAAATTTCTGAGCT 1164  
Db 181 ACAAGAAATTAAGCGCGAGGACTCATTTGCGGTTCTTCTGCTGTTTAAATTTCTGAGCT 240  
QY 1165 CTTTGTGTTTATCGCGCGGACCTTGGGGTGTGATCTTTTTCATGCGCGCGGATCGT 1224  
Db 241 CTTTGTGTTTATCGCGCGGACCTTGGGGTGTGATCTTTTTCATGCGCGCGGATCGT 300  
QY 1225 GTCGATATTATGCGTGGGTTGGCATCGCTTACCTGTTATGTTTGGCTTGGCTGCGAGC 1284  
Db 301 GTCGATATTATGCGTGGGTTGGCATCGCTTACCTGTTATGTTTGGCTTGGCTGCGAGC 360  
QY 1285 GAAAGACGCCATGACAAACCAAGGTGAAGCGCCACAGATCATTTGAAGAACAGAACCAAC 1344  
Db 361 GAAAGACGCCATGACAAACCAAGGTGAAGCGCCACAGATCATTTGAAGAACAGAACCAAC 420  
QY 1345 GTCGCGGATGACAGCGCTTTGGCGGTTTCGCGGTCGCTGACAGCGCGACCGGCT 1404  
Db 421 GTCGCGGATGACAGCGCTTTGGCGGTTTCGCGGTCGCTGACAGCGCGACCGGCT 480  
QY 1405 GCGGCTGGAGTGAAGCGTTCGATGAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCGCAT 1464  
Db 481 GCGGCTGGAGTGAAGCGTTCGATGAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCGCAT 540  
QY 1465 CCGTCTGACCTGTTGAACCCGATGCTTATGACGCGCTTTGCTTTATCGCGCGCT 1524  
Db 541 CCGTCTGACCTGTTGAACCCGATGCTTATGACGCGCTTTGCTTTATCGCGCGCT 600  
QY 1525 CCGCGCGATACGCGGACCGGACCGGATGCTTTTCCCGCTGGCGCTTCGCGGCAAG 1584  
Db 601 CCGCGCGATACGCGGACCGGACCGGATGCTTTTCCCGCTGGCGCTTCGCGGCAAG 660  
QY 1585 CCGATCTGCTTCCCGCTGGGTTTCGCGGACGAGCATGTGACGCGCGCTGTCAG 1644  
Db 661 CCGATCTGCTTCCCGCTGGGTTTCGCGGACGAGCATGTGACGCGCGCTGTCAG 720

QY	1645	CCCCAAGGTGTGGCGTGTGATCAACGTCGTCGTGGCAGTTGTGATGACCGCATTGGCCAT	1704
Db	721	CCCCAAGGTGTGGCGTGTGATCAACGTCGTCGTGGCAGTTGTGATGACCGCATTGGCCAT	780
QY	1705	CAAACCTGATGTTGATGGTGTAGTTTTCGCGGGTTTGGAAATC	1746
Db	781	CAAACCTGATGTTGATGGTGTAGTTTTCGCGGGTTTGGAAATC	822

  

RESULT 9	
ACC80941	
ID	ACC80941 standard; DNA; 711 BP.
XX	ACC80941;
XX	
XX	27-OCT-2003 (revised)
DT	11-AUG-2003 (first entry)
XX	LysE protein encoding sequence.
DE	
XX	L-lysine; L-arginine; LysE; ds.
KW	Corynebacterium glutamicum.
XX	
OS	
XX	
XX	Key Location/Qualifiers
PH	1..711
FT	/*tag= a
FT	/product= "lysE protein"
FT	
XX	
XX	EP1266366-A2.
PN	
XX	
XX	18-DEC-2002.
PD	
XX	
XX	05-JUN-2002; 2002EP-00012539.
PF	
XX	
XX	12-JUN-2001; 2001JP-00177075.
PR	
XX	
XX	(AJIN ) AJINOMOTO CO INC.
PA	
XX	

DR W11; 2005-251117-21.  
 DR P-PSDB; ABR58213.  
 XX Novel DNA encoding variant of LysE protein from a coryneform bacterium,  
 PT when introduced into methanol assimilating bacterium, facilitates  
 PT excretion of L-lysine and/or L-arginine to outside of a cell.  
 XX Example 1; Page 17-18; 23pp; English.  
 PS  
 XX The present invention relates to DNA encoding variants of protein with  
 CC loop region and six hydrophobic helices which facilitates excretion of L-  
 CC lysine and/or L-arginine to outside of cell of a methanol assimilating  
 CC bacterium when introduced into the bacterium. The method is used for  
 CC encoding a protein which facilitates excretion of L-lysine, L-arginine or  
 CC both of these L-amino acids to outside of a cell of a methanol  
 CC assimilating bacterium when DNA of the method is introduced into the  
 CC bacterium. The present sequence represents a lysE protein from  
 CC Bravibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003  
 CC to standardise OS field)  
 XX Sequence 711 BP; 135 A; 173 C; 222 G; 181 T; 0 U; 0 Other;  
 SQ  
 Query Match 29.9%; Score 711; DB 8; Length 711;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-211;  
 Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 OY 1016 ATGGTGATCATGGAAATCTTCATTACAGTCTGCTTTTGGGGGGCAGTCTTTTACTGTCC 1075  
 DB 1 ATGGTGATCATGGAAATCTTCATTACAGTCTGCTTTTGGGGGGCAGTCTTTTACTGTCC 60

Thu Mar 18 12:30:59 2004

1676 GTGGCAGTGTGATGACCGCATGTCCTCAAACTGATGTTGATGGGT 1723  
 661 GTGGCAGTGTGATGACCGCATGTCCTCAAACTGATGTTGATGGGT 708

XX WPI; 2001-376931/40.  
 DR P-PSDB; AAG93201.  
 XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analyzing  
 PT expression profile or pattern of a gene and identifying homologous gene.  
 XX Claim 1; SEQ ID NO 3455; 246pp + Sequence Listing; English.  
 PS The present invention provides a number of nucleotide and protein  
 XX sequences from the Corynebacterium bacterium Corynebacterium glutanicum. These  
 CC sequences are useful for identifying the mutation point of a gene derived from a  
 CC mutant of corynebacterium, measuring expression amount and analyzing  
 CC the expression profile or expression pattern of a gene derived from  
 CC Corynebacterium bacterium, and identifying a homologue of a gene derived from  
 CC Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino  
 CC acids, nucleic acids, vitamins, saccharides and organic acids.  
 CC particularly L-lysine. The present sequence is a nucleic acid described  
 CC in the exemplification of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from the European Patent Office  
 XX Sequence 708 BP; 134 A; 173 C; 221 G; 180 T; 0 U; 0 Other;  
 SQ

Query Match 29.8%; Score 708; DB 5; Length 708;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-210;  
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1016 ATGGTGTATCATGGAATCTTATTAAGTCTGCTTTGGGGGCGAGTCTTTTACTGTCC 1075  
 1 ATGGTGTATCATGGAATCTTATTAAGTCTGCTTTGGGGGCGAGTCTTTTACTGTCC 60  
 1076 ATCGACCGCGAGTACTGCTGATTAACAAAGGAATTAAGCGCGAAGACTCATTCGCG 1135  
 61 ATCGACCGCGAGTACTGCTGATTAACAAAGGAATTAAGCGCGAAGACTCATTCGCG 120  
 1136 GTTCTTCGCTGTTTAAATTTCTGACGCTCTTTTTCATCGCGGCGACTTGGCGCT 1195  
 121 GTTCTTCGCTGTTTAAATTTCTGACGCTCTTTTTCATCGCGGCGACTTGGCGCT 180  
 1196 GATCTTTTGTCCAAATCGCGCGCATGCTGCTGATTAATGCTGCTGGGTGCGATCGCT 1255  
 181 GATCTTTTGTCCAAATCGCGCGCATGCTGCTGATTAATGCTGCTGGGTGCGATCGCT 240  
 1256 TACTCTTATGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1315  
 241 TACTCTTATGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 1316 CCACAGATCATTTGAAGAAACAGAAACCAACCGTGGCGGATGACACCGCTTTGGCGGTTCG 1375  
 301 CCACAGATCATTTGAAGAAACAGAAACCAACCGTGGCGGATGACACCGCTTTGGCGGTTCG 360  
 1376 GCGTGGCGCTGACACCGCGCAACCGGCTGGCGGATGAGCGTGCATTAAGCAGCGG 1435  
 361 GCGTGGCGCTGACACCGCGCAACCGGCTGGCGGATGAGCGTGCATTAAGCAGCGG 420  
 1436 GTTGGGTGAAGCCCATGTTGATGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1495  
 421 GTTGGGTGAAGCCCATGTTGATGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 1496 TTGGACCGCTTGTGTTTATCGGCGGCTGCGCGGCAATACCGCGACCGCGGCTGG 1555  
 481 TTGGACCGCTTGTGTTTATCGGCGGCTGCGCGGCAATACCGCGACCGCGGCTGG 540  
 1556 ATTTTCGCGCTGCGCGGTTTCGCGCAAGCTCATCTGCTTCCCGCTGGTGGGTTCGCG 1615  
 541 ATTTTCGCGCTGCGCGGTTTCGCGCAAGCTCATCTGCTTCCCGCTGGTGGGTTCGCG 600  
 1616 GCAGCAGCATTTGACGCGCGCTGTCCAGCGCCCAAGGTGTGGCGTGGATCAACGCTGC 1675  
 601 GCAGCAGCATTTGACGCGCGCTGTCCAGCGCCCAAGGTGTGGCGTGGATCAACGCTGC 660

RESULT 11

AAH45375

ID AAH45375 standard; DNA; 1568 BP.

XX AAH45375;

AC AAH45375;

XX 11-SEP-2001 (first entry)

DT C. thermaminogenes lysin biosynthetic enzyme lyse DNA.

XX Heat-resistant; lysin biosynthesis; enzyme; coryneform;

XX aspartate-semialdehyde dehydrogenase; lyse; ds.

XX Corynebacterium thermoaminogenes.

XX JP2001120270-A.

XX 08-MAY-2001.

PD 01-NOV-1999; 99JP-00311148.

XX 01-NOV-1999; 99JP-00311148.

XX (AJIN ) AJINOMOTO KK.

XX WPI; 2001-364760/38.

XX P-PSDB; AAG64047.

XX A heat-resistant lysin biosynthetic system enzyme gene of a high

XX temperature-resistant coryneform microbe.

XX Example 5; Page 22-24; 27pp; Japanese.

XX The invention relates to a gene from a high temperature-resistant

XX coryneform microbe that encodes a heat-resistant lysin biosynthetic

XX enzyme. The enzyme has aspartate-semialdehyde dehydrogenase activity and

XX can be used for growing amino acid-producing microbes. The present

XX sequence encodes an enzyme of the invention

XX Sequence 1568 BP; 258 A; 525 C; 490 G; 295 T; 0 U; 0 Other;

SQ

Query Match 29.4%; Score 698.6; DB 4; Length 1568;

Best Local Similarity 68.4%; Pred. No. 7e-207;

Matches 1017; Conservative 0; Mismatches 454; Indels 16; Gaps 3;

240 GGGTTTCGGGAGAGTCCCAACCAAGGCTTCGGCGAATTTGCTTCAAAACCTTCG 299

1 GGGCTTCGGGAGAGGCCCCCAAGGCTTCGGCGAATTTGCTTCAAAACCTTCG 60

300 CCGACGCGACATGATGATACGGCGCTTCGGCGGATTTGCTTCAAAACCTTCG 359

61 CAGACGCGACATGATGATACGGCGCTTCGGCGGATTTGCTTCAAAACCTTCG 120

360 CACGGTCTTGAAGCAATCTTTGGGACCGGAGCTTAAGACGGGAT--CGCAGCCCAATCT 418

121 CGCGTCTTCGACGACATCATTTGGGGCGGAAACGCGATGACCGGATGCGCAGCTG 180

419 AGTTTCCCATCAACCATGTAGGCAATCCCGCAATGAGGGGGTTGCAATGGCGAGTGGCG 478

181 GGTTCGGCATCCAGGTGTACCGTCTTCAGTTCCGGGGTGGCGACCGGAGGTGTCTC 240

479 ATGGTTTCAAGTCTTACTACTTTCACATCCCGCAGCGGATTTAGCTTACGGGTTCAG 538

241 ATGACCCCGAGGCGTAGGACCTTCAACCCGCGCAAGGGGTTCGGCTTCGGGGTTCAG 300

539 CCTTAAACATCTTCAACCGCGGAGGAGGATTAATGTGTGGCTTTCATCTTCCAGGCGAG 598

301 CCGACGACGGAACCGCGGCGCAACGAGGAGTGTGGGCTTCTCTTCCACACGCGAGC 360

1679 GCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTA 1725  
1426 GCGTGTGTGCTCACCAGATTGGCGTGAAGCTCATCTGATGGGTTA 1472

RESULT 12  
ACC80942  
ID ACC80942 standard; DNA; 712 BP.

XX ACC80942;  
XX 27-OCT-2003 (revised)  
DT 11-AUG-2003 (first entry)

XX LysE24 protein encoding sequence.  
XX L-lysine, L-arginine; LysE24; ds.  
XX Corynebacterium glutamicum.

XX Key Location/Qualifiers  
CDS 1..711  
/\*tag= a  
/product= "lysE protein"

PN BP1266966-A2.  
XX 18-DEC-2002.  
XX 05-JUN-2002; 2002EP-00012539.  
XX 12-JUN-2001; 2001JP-00177075.  
XX (AJIN ) AJINOMOTO CO INC.  
XX Gunji Y, Yasueda H;  
XX WPI; 2003-241171/24.  
XX P-PSDB; ABR58214.

Novel DNA encoding variant of LysE protein from a coryneform bacterium, when introduced into methanol assimilating bacterium, facilitates excretion of L-lysine and/or L-arginine to outside of a cell.

Example 1; Page 19-20; 23pp; English.

The present invention relates to DNA encoding variants of protein with loop region and six hydrophobic helices which facilitates excretion of L-lysine and/or L-arginine to outside of cell of a methanol assimilating bacterium when introduced into the bacterium. The method is used for encoding a protein which facilitates excretion of L-lysine, L-arginine or both of these L-amino acids to outside of a cell of a methanol assimilating bacterium when DNA of the method is introduced into the bacterium. The present sequence represents a lysE24 protein from CC Brevibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003 to standardise OS field)

Sequence 712 BP; 133 A; 173 C; 224 G; 182 T; 0 U; 0 Other;  
Query Match 29.4%; Score 696.8; DB 8; Length 712;  
Best Local Similarity 99.6%; Pred. No. 1.6e-206;  
Matches 709; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1016 ATGTGATCATGGAATCTTCAATTACAGGTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 1075  
1 ATGTGATCATGGAATCTTCAATTACAGGTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 60

1076 ATCGACCGCGAGATGCTACTGTGATTAACAGGAATTAACCGGAGGACTCAATTGCG 1135  
61 ATCGACCGCGAGATGCTACTGTGATTAACAGGAATTAACCGGAGGACTCAATTGCG 120

1136 GTTCTTCTCGTGTCTTAAATTTCTGAGCTCTTTTCTTCATCGCGGACCTTGGCGGT 1195

599 GTGAGCGTTGCTCCACCCCAAGAGCTTACTCTGTGTAACACGCGAGGAAACCATGTGGAT 658  
361 GTGAGGTGACCGCACCCCAATGTGCAACCTCGCGAACAACGCGCGGAAACCGAGTGGAC 420

659 AGCGAATCTCGTTGATGCGATGTTTACGGGATTTTCAGCAAGGCTCCAGATAGTTGC 718  
421 AGGGAATCGCGGTGATGACCAACCGTTCAGGGGATCTCTGTCAGCCGTTTCGCCAGTTGC 480

719 GCTTTAGTTCTGCTTTCAGCAACACCATTTTCGCGCTGCTTTCACAGGACTTTCACCC 778  
481 TCAGGGTCTCCGCTTCGACAGCGCCATCTTTGCGGCGCGCTTGACACGACACCTCCCG 540

779 GCTTCGGTGTCTTTGCGCGGTGCGGATACCAACTGACCCGATGATGCTGCG 838  
541 GCTTCGGTGGCAACCGCGGTGCGGATACCAACTGACCCGATGATGCTGCG 600

839 AGAGCTTTACGGCGCTGACTACCGCGGAGGAAATGGAAGGCTAAGAGCGCGCT 898  
601 AGTGCCTTGTAGCGCTGCTGACCGCGGAGGAGATGACAGTGGAGGAGCGGTC 660

899 TGAAGCTGCTTCATCAATGATTCGAGGCAAGTGTCCAGTGAATGGGTTTCATGAAG 958  
661 TGAAGCTGCTTCATCAATGATTCGAGGCAAGTGTCCAGTGAATGGGTTTCATGAAG 720

959 CTATATTAAACCATGTTAAGAACCAATCTTTTACTTAACTTCTCATAGTTCAGATG 1018  
721 CAATTTAAACCATGTTCACTCTCTATCAATTTTACTTAACTTCTCATAGTTCAGATG 780

1019 GTGATCATGGAATCTTCATACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTCCATC 1078  
781 CGGACATGGAATCTTCATACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTCCATC 840

1079 GGAACGCAAGATGATCTGTGATTAACAGGAATTAAGCGGAGGACTCATTTGGGTT 1138  
841 GGCCACAGATGCTCTGTGATCAACAGGCAATCAACAGGAGGCAATCAACGCGCTC 900

1139 CTCTCTCGTGTGTTAAATTTCTGAGTCTTTTGTTCATCGCGCACTTTGGCGGTGAT 1198  
901 ATCATGCTGTCTGCTGCTCCGAGCTGCTGTTTCAACCTCGGACCTTCGGGTCGCG 960

1199 CTTTTGTCAATGCGCGCGATCGTGTGATATATGCTGCGGTGCGGCTCGCTTAC 1258  
961 CTGATCTCGACACGCGCGGATCATCTCGACATCTCGCTGTGTGCGGATCGCTTAC 1020

1259 CTGTTATGTTTTCGCTCATGCGAGCAAGACGCCATGACAAACAAGGTGGAAGCGCA 1318  
1021 CTGCTGTGTTGCGGTGATGCGCGGCGCGAGCGCTGCGCGCGCGCACCGAGTAACC 1080

1319 CAGATCATTTGAAGAAACAGAACCAACCGTGGCGATGACACGCTTTGGGCGTTGCGG 1378  
1081 TTTGTC-----GAGCATTCGAAACCGGTTG---CGGACGCTGCGCTCGCG 1125

1379 GTGCGCACTGACACGCGCAACCGGCTGCGGTGGAGGTGAGCTGATGAAGCAGCGGTT 1438  
1126 GGGGCGGTGACGACGAACCAACGACCGCGCTCCGCATCATCAGGACACCGCGAGTCC 1185

1439 TGGGTAAAGCCATGTTGATGGCAATCTGCTGACCTGTTGTAACCCGATGCGTATTG 1498  
1186 TGGGTGAGGCGCATGCTCATGCGCATTTGCTGACCTGGCTCAATCCCAATGCCCTAC 1245

1499 GAOCGGTGTGTGTTATGCGCGGTGCGGCGGATACGCGGACACCGGACCGGTGATT 1558  
1246 GATGCTTCTCTTCAATCGTGTGTGCGAGCCGATACGCGGAGACCGGTGCGTATC 1305

1559 TTGCGCGCTGCGCGGTTCGCGCAAGCCTGATCTGCTTCCGCTGCTGCTGTTTTCGGGCG 1618  
1306 TTCGTGCGGCGCTTTCGCGCGGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1365

1619 GCAGCATTTGTCAGCGCGCTGTCCAGCGCCCAAGGTGTGCGCTGCTGATCAACGCTGCTG 1678  
1366 GCGGACATGTCGCTCCCTGTCAGCGCGCGGTCTGCGCTGCTGATCAACATAGGTGTG 1425







Thu Mar 18 12:30:59 2004

AC AAF68077;  
XX 11-APR-2001 (first entry)  
XX  
XX Corynebacterium glutamicum MCT protein encoding DNA SEQ ID NO:669.  
XX  
XX Corynebacterium glutamicum; brevivibacterium lactofermentum; MCT;  
XX membrane construction and membrane transport protein; petroleum spill;  
XX hydrocarbon degradation; gram positive aerobic bacterium; marker;  
XX identification; microorganism; fine chemical production; transformation;  
XX genome mapping; genetic engineering; ds.  
XX  
XX Corynebacterium glutamicum.  
XX  
XX WO200100805-A2.  
XX  
XX 04-JAN-2001.  
XX  
XX 23-JUN-2000; 2000WO-IB000926.  
XX  
XX 25-JUN-1999; 99US-0141031P.  
XX 08-JUL-1999; 99DE-01031454.  
XX 08-JUL-1999; 99DE-01031478.  
XX 08-JUL-1999; 99DE-01031563.  
XX 09-JUL-1999; 99DE-01032122.  
XX 09-JUL-1999; 99DE-01032124.  
XX 09-JUL-1999; 99DE-01032125.  
XX 09-JUL-1999; 99DE-01032128.  
XX 09-JUL-1999; 99DE-01032180.  
XX 09-JUL-1999; 99DE-01032182.  
XX 09-JUL-1999; 99DE-01032190.  
XX 09-JUL-1999; 99DE-01032191.  
XX 09-JUL-1999; 99DE-01032209.  
XX 09-JUL-1999; 99DE-01032212.  
XX 09-JUL-1999; 99DE-01032227.  
XX 09-JUL-1999; 99DE-01032228.  
XX 09-JUL-1999; 99DE-01032229.  
XX 09-JUL-1999; 99DE-01032230.  
XX 14-JUL-1999; 99DE-01032927.  
XX 14-JUL-1999; 99DE-01033005.  
XX 14-JUL-1999; 99DE-01033006.  
XX 27-AUG-1999; 99DE-01040764.  
XX 27-AUG-1999; 99DE-01040765.  
XX 27-AUG-1999; 99DE-01040766.  
XX 27-AUG-1999; 99DE-01040830.  
XX 27-AUG-1999; 99DE-01040831.  
XX 27-AUG-1999; 99DE-01040832.  
XX 27-AUG-1999; 99DE-01040833.  
XX 31-AUG-1999; 99DE-01041378.  
XX 31-AUG-1999; 99DE-01041379.  
XX 31-AUG-1999; 99DE-01041395.  
XX 03-SEP-1999; 99DE-01042077.  
XX 03-SEP-1999; 99DE-01042078.  
XX 03-SEP-1999; 99DE-01042079.  
XX 03-SEP-1999; 99DE-01042088.  
XX  
XX (BADI ) BASF AG.  
XX  
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
XX  
XX WPI; 2001-071486/08.  
XX P-PSDB; AAB76844.  
XX  
XX Corynebacterium glutamicum nucleic acids encoding membrane construction  
XX and membrane transport proteins or their portions, useful for typing or  
XX identifying C. glutamicum or related bacteria, and as markers for  
XX transformation.  
XX  
XX Claim 3; Page 1107-1108; 1119pp; English.  
XX  
XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane  
XX construction and membrane transport (MCT) proteins given in AAB76510 to  
XX AAB76847. The MCT nucleic acids and proteins are useful in the

CC identification of microorganisms which can be used to produce fine  
CC chemicals, for modulating fine chemical production in C. glutamicum or  
CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or  
CC identification of C. glutamicum or related bacteria, as reference points  
CC for mapping C. glutamicum genome, and as markers for transformation.  
CC AAF68082 and AAF68082 represent sequencing primers which are used in an  
CC example from the present invention  
XX  
XX Sequence 993 BP; 222 A; 247 C; 277 G; 247 T; 0 U; 0 Other;  
SQ  
Query Match 20.4%; Score 485.2; DB 4; Length 993;  
Best Local Similarity 99.0%; Pred. No. 2.7e-140;  
Matches 499; Conservative 0; Mismatches 3; Indels 2; Gaps 1;  
1873 AGCACCAATCAATGCACTGCTACGGTA--TCGCGCGGTACTCTCTTGTCTCGGCGAGC 1930  
993 AGCACCAATCAATGCACTGCTACGGTAATTCGCGCGGTACTCTCTTGTCTCGGCGAGC 934  
1931 ACCCATGCAAGCGCCATCTGCGCAAGTGACTGCGCGGTACTCTCTTGTCTCGGCGAGC 1990  
933 ACCCATGCAAGCGCCATCTGCGCAAGTGACTGCGCGGTACTCTCTTGTCTCGGCGAGC 874  
1991 TTGCGGACCATATCAATATTTGTTCACTGCTTCACTGCTTCACTGCTTCACTGCTTCACTGCT 2050  
873 TTGCGGACCATATCAATATTTGTTCACTGCTTCACTGCTTCACTGCTTCACTGCTTCACTGCT 814  
2051 CTGCGCGGGGAAACCTCTCGAATTCATTCGAGATATTTGTTGCGGAGGAGGAGGAGGAGGAGG 2110  
813 CTGCGCGGGGAAACCTCTCGAATTCATTCGAGATATTTGTTGCGGAGGAGGAGGAGGAGGAGG 754  
2111 AGTGTGTGAAGCAATGAGCAAGCAAGCAATTTGTTGCGGAGGAGGAGGAGGAGGAGGAGGAGG 2170  
753 AGTGTGTGAAGCAATGAGCAAGCAAGCAATTTGTTGCGGAGGAGGAGGAGGAGGAGGAGGAGG 694  
2171 CCTCATCTGCGCGGTCTCTCCACCAACGATTAATGATGAATGATGATGATGATGATGATGATGAT 2230  
693 CCTCATCTGCGCGGTCTCTCCACCAACGATTAATGATGAATGATGATGATGATGATGATGATGAT 634  
2231 AGAAGCGGAGCGCTCTCTCGCGCATGAACTCAGCGCGCTCTCGCGCTCTCGCGCTCTCGCGCT 2290  
633 AGAAGCGGAGCGCTCTCTCGCGCATGAACTCAGCGCGCTCTCGCGCTCTCGCGCTCTCGCGCT 574  
2291 TAGGAGGAATACCCAGTAAGAGCGCTTTCCAGACGCAATGTCAGCAATGTCAGTAC 2350  
573 TAGGAGGAATACCCAGTAAGAGCGCTTTCCAGACGCAATGTCAGCAATGTCAGTAC 514  
2351 ATGTTTCTTCCAAAAGGAGTATCT 2374  
513 ATGTTTCTTCCAAAAGGAGTATCT 490

RESULT 15  
ACA29651/c  
ID ACA29651 standard; DNA; 1095 BP.  
XX  
XX ACA29651;  
XX  
XX 19-JUN-2003 (first entry)  
XX  
XX Prokaryotic essential gene #11308.  
XX  
XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
XX drug design; gene.  
XX  
XX Corynebacterium diphtheriae.  
XX  
XX WO200277183-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 21-MAR-2002; 2002WO-US009107.  
XX  
XX 21-MAR-2001; 2001US-00815242.  
XX

us-09-105-117k-1.rng

Thu Mar 18 12:30:59 2004

1970 TCCTGGCGATGTCAATGAGCTTGGGACCATATCAATATTTGTTTCAAGTCCATGCC 2029  
 854 CGTAAAGCTTAAATCATTCAGTGCACGACCATGTCAGATTTTAGCGTTGAGCTCTCT 795  
 2030 TCAGACAGGAGACTTACCTGGCTGGCGCGGAAACCTCTTGGAAATTCATCGAGATATTG 2089  
 794 TTGGCCAAAGGACTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 735  
 2090 TCCGTGAGCAGGCGCTCGCAGCAAGTGGTGAAGAAATGACGCAAGCAATTTGTTGCA 2149  
 734 TCAGTTAAAGCCCTGTGCAAGTGGCGAAATGCAATAACACCACTCCGTTATTTCGCC 675  
 2150 GTGACTGGAACAGATTTCTACCGTTCATCGCGCGTTCCTCCACCCCAACGATTAATGATG 2209  
 674 GCAGACTCTAAACAGTTTCGCGCATCTTCACCTGGTTCTTCTACCCAGCGTTCAAGATA 615  
 2210 GAATAGCTTGGCTGATGATCAATGAGAGCGGCGAGCCCTCTCCGCCATGACTCAGCGGCC 2269  
 614 GAATAGCTCGGTTGATGATTCAGCAGCGGCGAGCCCTCACCAGCTAGAAATTTCCGAGATC 555  
 2270 TCCGCTGTGAGCTCTGGACCGTAGGAGAAATACCGTAAAGAGCCTTTCCAGACGCA 2329  
 554 TCTGCGGTTAGTTCAAGTTCATAGGAGGAAATTTCCCATAGGAGCCTTCCCGATGCC 495  
 2330 ACAATGTCAGCAATGCGTACATGTTTCTTCCAAAGAGATCT 2374  
 494 ACGATGTCTCGCAAAAGCATATGCGAGTTTCTTCCAAAGCGGTATCT 450

Search completed: March 15, 2004, 14:59:10  
 Job time : 974.046 secs

PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI: 2003-029926/02.  
 DR P-PSDB; ABU25781.  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX Claim 14; SEQ ID NO 17521; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway; (8)  
 CC required for proliferation, or that inhibits cellular proliferation; (9)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC not form part of the printed specification. Note: the sequence data for this patent did  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX  
 XX Sequence 1095 BP; 264 A; 247 C; 311 G; 273 T; 0 U; 0 Other;  
 SQ  
 Query Match 12.9%; Score 305.8; DB 7; Length 1095;  
 Best Local Similarity 67.1%; Pred. No. 3.5e-84;  
 Matches 433; Conservative 0; Mismatches 212; Indels 0; Gaps 0;  
 1730 TCGCGGGTTTGGAAATCGGTGGCTTCGCCCAAAATGTTGATGCGCGCGTCTGGGAAATC 1789  
 1094 TCGTGAACCTTCGAGCGAGTAGCACCAGCCGCGCCCAATATTAATTCAGCGTCTTGGCAACA 1035  
 1790 TCATCGATCGCTCCAACTCGCGGTGAGAAATCTCCAAAGTTGTTGATGAAATCAGGCTG 1849  
 1034 GTGTCAATATATTTGCGTTCTTCAACGCTGAATCAAGATTTATAGAGCTCCAGATTT 975  
 1850 TTGTCCAGCTGTCTCACTGACGAAGACCAATCAATGCACTGTGTACGTTATCCGCGCG 1909  
 974 TGATCCAACTGCGCTACCGAGATGCTCCGATCAACGACCTGTGTACCGTTGTGCCCCCA 915  
 1910 TACTCTCTTGTCTCGCGAGCAGCCATGCAAGCGCATCTGCGCAAGTCACTGCGCGCT 1969  
 914 TAATCGCTTGTTCGCGAGCAGCCAGCCGCAATAGCAGCTGTGCGAACGTTTGGCCAGT 855

## ALIGNMENTS

FLING DATE: 17-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunter, Tom  
REGISTRATION NUMBER: 38,498  
REFERENCE/DOCKET NUMBER: 15371A-17  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/543/9600  
TELEFAX: 415/543/5043  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15239 base pairs  
TYPE: nucleic acid  
STRANDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-390-878-17

File	Score	DB 1	Length
15239	122	8	15239

Query Match	5.2%	Score 122.8;	DB 1;	Length 15239;
Best Local Similarity	49.7%	Pred. NO. 1.le-25;		
		Mismatches 422;		
		Indels 18;		
		Gaps 4;		

78 GTAACTAAGGGCCGCAATCCCTCGATTGTGTCATCAACGACGGCGTCTGTGAGTCTAGTCA 137  
6117 GTTCTGGGCCCCGGTACAGACCGCTTGGCGGGCCCTACCGTGTGGTATTTCGGCGCA 6058

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	122.8	5.2	15239	1	US-08-390-878-17	Sequence 17, Appl
C 2	121.2	5.1	909	4	US-09-894-844-12	Sequence 12, Appl
C 3	121.2	5.1	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C 4	121.2	5.1	4411529	3	US-09-103-840A-1	Sequence 1, Appl
C 5	100.8	4.2	5541	1	US-08-920-812-20	Sequence 20, Appl
C 6	100.8	4.2	5541	1	US-08-920-827-20	Sequence 20, Appl
C 7	100.8	4.2	5541	1	US-08-921-177-20	Sequence 20, Appl
C 8	100.8	4.2	5541	1	US-08-362-577C-20	Sequence 20, Appl
C 9	100.8	4.2	5541	2	US-08-920-828-20	Sequence 20, Appl
C 10	81.4	3.4	1095	4	US-09-489-039A-5370	Sequence 5370, Ap
C 11	80.2	3.4	915	4	US-09-489-039A-4735	Sequence 4735, Ap
C 12	73.4	3.1	936	4	US-09-543-681A-791	Sequence 791, App
C 13	65.2	2.7	900	4	US-09-252-991A-7644	Sequence 7644, Ap
C 14	64	2.7	1002	4	US-09-724-623-24	Sequence 24, Appl
C 15	61.8	2.6	699	4	US-09-252-991A-7908	Sequence 7908, Ap
C 16	61.8	2.6	834	4	US-09-252-991A-7643	Sequence 7643, Ap
C 17	61.8	2.6	894	4	US-09-252-991A-7841	Sequence 7841, Ap
C 18	57	2.4	597	4	US-09-894-844-11	Sequence 11, Appl
C 19	56.6	2.4	642	4	US-09-489-039A-4674	Sequence 4674, Ap
C 20	54.6	2.3	903	4	US-09-328-352-1694	Sequence 1694, Ap
C 21	53.8	2.3	1176	4	US-09-252-991A-7571	Sequence 7571, Ap
C 22	49.2	2.1	921	4	US-09-252-991A-14393	Sequence 14393, Ap
C 23	49.2	2.1	1905	4	US-09-252-991A-14312	Sequence 14312, A
C 24	47.8	2.0	474	4	US-09-252-991A-5488	Sequence 5488, Ap
C 25	46.2	1.9	903	4	US-09-252-991A-7723	Sequence 7723, Ap
C 26	45.8	1.9	720	4	US-09-252-991A-7723	Sequence 7723, Ap
C 27	45.4	1.9	45613	4	US-09-596-002-22	Sequence 22, Appl

Thu Mar 18 12:30:59 2004

PRIOR APPLICATION NUMBER: 60/097,936  
 PRIOR FILING DATE: 1998-08-25  
 NUMBER OF SEQ ID NOS: 137  
 SOFTWARE: EMBLSEQ for Windows Version 4.0  
 SEQ ID NO 12  
 LENGTH: 909  
 TYPE: DNA  
 ORGANISM: Mycobacteria tuberculosis  
 US-09-894-844-12

Query Match 5.1%; Score 121.2; DB 4; Length 909;  
 Best Local Similarity 49.5%; Pred. No. 5.9e-26;  
 Matches 433; Conservative 0; Mismatches 423; Indels 18; Gaps 4;

78 GTAACTAAGCGCGCAATCCCTCGATTGCTGTCATCAACGAGCGGCTGCTGTAGTCTAGCTA 137  
 Db  
 892 GTTCTGCGCGCGGTACAGACCGCTTGCCTGCGCGCGCTTCTCGAGGTGTATGTCG 833  
 QY  
 138 GAGATCTAGATTCAGAGCGCGCATCTGTTGCAATATCATCGGTGTCAATGCGGTATCTCAT 197  
 Db  
 832 TGATCGGACTGTCCAGTTTCCAGCATTTGCCAATAGAGAGGACGTCGAGGTGTATGTCG 773  
 QY  
 198 CGAGGAGGATCACTTCTCTCTCTCTTTTAGCATGAGGAGCAGCTTGGGTTTCGGGAAGAGTC 257  
 Db  
 772 AGACCCGTTACGACGATCCATCGGCAAGCGAGATGCTGCCAGGTTCTCGGGGAACATGC 713  
 QY  
 258 CCCAACCAAGGCTCGGCGAATTCCTCACCACCAACCTTCCGCGACGAGCAATGGATA 317  
 Db  
 712 CCCATCCGACCGCGCGCGCTGCGCGGTGAGCCCTCTGTGTCGGGACAAAGTCCG 653  
 QY  
 318 CGCGCTCGCGCCCAACAGGACCATCGAC---GCGCGCTCCAGGTACCGTCTTCAAGCA 374  
 Db  
 652 TCGGCTCTGGTGATGGCGGCGAGAGGCTTACGCAACCAATGTCTCGAGCCCATCGT 593  
 QY  
 375 CATCTTTGGGACCGAAGCGGTAGACGCGGATTCGATGCGCAAGTGGCGCATGTTTCAAGTTCTA 494  
 Db  
 592 CAGGATCCACGCGAGTACGAGCTTTAGCGCGCGCGGCGAGTGACCCCTCGATA 533  
 QY  
 435 TGATGCGATCCCGCAATGAGGGGTTGCAATGCGCAAGTGGCGCATGTTTCAAGTTCTA 494  
 Db  
 532 GATGCGCTGGAACAAATGCGCTGTGCGCACTGTGAGGTAGCGCATTTTCAACCGCGGT 473  
 QY  
 495 CTACTTCACTCCCGCAACGAGGATTTACCGGGTTACCGGTTACCGTCTTCAACATCTCCAC 554  
 Db  
 472 GCACCGCGGACGCCCGGCGACCGGTTCCGCTCGGTGTCACCGCGCCCATCGCCACACCT 413  
 QY  
 555 GCGCGACGAGGAATAATGTGTGGCTTCACTTCCAGGCGGAGCGGTGCTTCCAC 614  
 Db  
 412 CCGGTAGCAGCGCGCGGAAATGGTCTCTCGATCCGAAAGCTGCGAGGAGCGTCCG 353  
 QY  
 615 CCCAAGNAGCTACTCTGTTGACACCGGAGGAAACCATGTGATAGTAGCGAATCTCGGTGA 674  
 Db  
 352 -----CGAGACCGTGCAGACCGGCGGAACCATGTGCGCAATGGAATCGCGTTTA 302  
 QY  
 675 TGGCGATGGTTAAACGGGA---TTTCAGAGAGGT---CCAGATAGTTCGCTTTAGTTT 728  
 Db  
 301 CGGCAATGGTGTGCTGCGCTGCGTTTTCAGGCAAGCGTTTGCACCCCATTTTCAGGAGCGCT 242  
 QY  
 729 CTGCTTGAGGACACACCATTTTTCGCGCTGCTTGCACAAAGGACTTCCACCGCTTTCGGTTG 788  
 Db  
 241 CGGACTCGAGCAACGCTGTTTGGCGCGGCGCAACCGCAACAGGGGATACCTGCGGTCTGCG 182  
 QY  
 789 CTTTGGCGGTTGGGTGGCGGATACCAACACTTCGACCCCATGATGCTCGAGAGCTTAA 848  
 Db  
 181 CCGGACATGGCTTTTCCCTTGACCACTGACGACCTGCGCGCTGCTCTCCACGACTTGA 122  
 QY  
 849 CGCGCTGACTCACCGCGGAGGGGAAATGGAAGGGCTAGGAGGCGCTTCCGAGCTGC 908  
 Db  
 121 TCGGCTGACTGACGAGCGGAGGGGTGACATGAGGCGCTCCGCGCGCGCATCGAAGCTGC 62  
 QY  
 909 CTTTCATCAATGATGAGACCAAGTGTCCAGTTG 942  
 Db  
 61 CCAAGTTCCAGCAGCGGACCAATGCGGCCAGCTG 28

138 GAGATCTAGATTCAGAGCGCGCATCTGTTGCAATATCATCGGTGTCAATGCGGTATCTCAT 197  
 Db  
 6057 TGNCTGGACTGTCCAGTTTCCAGCATTTGCCAATAGAGAGGACGTCGAGGTGTATGTCGCG 5998  
 QY  
 198 CGAGGAGGATCACTTCTCTCTCTTTAGCATGAGGAGCAGCTTGGGTTTCGGGAAGAGTTC 257  
 Db  
 5997 AGACCCGTTACGACGATCCATCGGCAAGCGGAGATGCTGCGAGCTTCTCGGGGAACATGC 5938  
 QY  
 258 CCCAACCAAGGCTCGGCGAATTTGCTTCCACCAACCTTCCGCGACGAGCAATGGATA 317  
 Db  
 5937 CCCATCCGACCGCGCGGCTGCGCGGTGAGCCCTCTGTGTCGGGACCAAGTCCG 5878  
 QY  
 318 CGCGCTCGCGCCCAACAGGACCATCGAC---GCGCGCTCCAGGTCAGGCTTGAAGCA 374  
 Db  
 5877 TCGGCTGTGGTATGCGCGACGAAAGGCTTACGCAACATGCTGTCGAGCCCATCGT 5818  
 QY  
 375 CATCTTTGGGACCGAGGATTAAGACGCGCATCGAGCCCAATCTAGTTTCCCATCAACCA 434  
 Db  
 5817 CAGGATTCACACGCGAGTGAAGAGCTTTAGCGCGCGGCGGAGTGAACCGTTCGATA 5758  
 QY  
 435 TGTAAGCATCCCGCAATGAGGGGTTGCAATGGCCAAAGTGGCGCATGCTTCCAGTTCTA 494  
 Db  
 5757 GATGCGCTGACGATGGCTGCTGCGCACTGTGTAGTAGCGCATTTTCAACCGAGGT 5698  
 QY  
 495 CTACTTCACTCCCGGACGAGGATTAAGTCTTACGCGGTACCGCTTCCAAACATCTCCAC 554  
 Db  
 5697 GCACCCGCGACCGCGGACCGGCTTCCAGGCTGCGGTGCTTCCAGGCGGAGTGCCTCCAC 614  
 QY  
 555 GCGCGACGAGGAATAATGTGTGGCTTCACTTCCAGGCGGAGGCTGAGCGTGTGCTCCAC 674  
 Db  
 5637 CCGGTAGGAGCGCGCGGAGATGCTTCTGCTGCTGCTGCTCGATCCGAAAGCTGAGGAGCTGC 5578  
 QY  
 615 CCCAAGAGCTTACTCTGTTGAACACGCGGAGGAAACCATGTGATGAGGATCTCGGTGA 674  
 Db  
 5577 -----CGAGACCGTGCAGACCGCGCGGAAACCATGTGCGCATGGAATCGGGTTTA 5527  
 QY  
 675 TGGCGATGGTTAAGGGA---TTTCAGAGGCGT---CCAGATAGTTCGCTTTAGTTT 728  
 Db  
 5526 CCGCAATGGTATCCGCGTGGCTTTTACGCGACCGGTTGCGACCATTTTCAGCGAGCGCT 5467  
 QY  
 729 CTGCTTGACGACACCATTTTCCGCGCTGCTTGCACAAAGGACTTCCACCGCTTGGTTG 788  
 Db  
 5466 CGGACTCGAGCAACGCTGTTTGGCGGCGCAACCGCAACAGGCGGATACCTGCGGTGCTG 5407  
 QY  
 789 CTTTGGCGGTTGGGTGCGGATACCAACACTGACCCGACGATGCTGCGAGGACTTAA 848  
 Db  
 5406 CCGGACATGGCTTTTCCCTGACCAACGACGACCTGGCGGACCTGCTGCTTCCAGACTTGA 5347  
 QY  
 849 CGCGCTGACTCACCGCGGAGGGAAATGGAAGGGCTAAGGAGGCGCTTCCAGCTGC 908  
 Db  
 5346 TGGGCTGACTGACGAGCGGAGGGGTGACATGATGAGCGCTCCGCGCGCGCATCGAAGCTGC 5287  
 QY  
 909 CTTTCATCAATGATGAGACCAAGTGTCCAGTTG 942  
 Db  
 5286 CCAAGTTCCAGCAGCGGACCAATGCGGCCAGCTG 5253

RESULT 2  
 US-09-894-844-12/c  
 ; Sequence 12, Application US/09894844  
 ; Patent No. 6686166  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Behr, Marcel  
 ; APPLICANT: Small, Peter  
 ; APPLICANT: Schoolnik, Gary  
 ; APPLICANT: Wilson, Michael A.  
 ; TITLE OF INVENTION: Molecular Differences Between Species of  
 ; the M. Tuberculosis Complex  
 ; FILE REFERENCE: STAN102CON  
 ; CURRENT APPLICATION NUMBER: US/09/894,844  
 ; CURRENT FILING DATE: 2001-06-27  
 ; PRIOR APPLICATION NUMBER: 09/318,191  
 ; PRIOR FILING DATE: 1999-05-25

Db 2226899 CCGCAATGGTGTATCCGGCTGCTTTTCAGGACGCGTTGCCACCCATTTTCAGGAGCGCT 2226958  
 Qy 729 CTGCTTCGACCAACACCAATTTTCGGCGCTGCTTTCAGGACGCGTTGCCACCCATTTTCAGGAGCGCT 788  
 Db 2226959 CGGACTCGACCAACGCTGTTTTCGGCGCAACCGCAACGAGCGGATACCTTCGGCTGCTGCG 2227018  
 Qy 789 CTTTGGCGGCTTGGGTGCGATACCAACCACTTCGACCGACGCTGATGCTCGAGAGCTTTAA 848  
 Db 2227019 CCGACATGCTTTTTCCTGACCAACGAGACCTTGGCCGACCTGCTCTCAACGACTTGA 2227078  
 Qy 849 CCGCTGACTCACCGCGGAGGCGGAATGGAAGGCTTAAGAGGCGCTTTCAGAGCTGC 908  
 Db 2227079 TCGCTGACTGACGAGCGGCGGCTGACATGAGGCGCTTCGCGCGGCGATCGAAGCTGC 2227138  
 Qy 909 CTTTCAATCAATGATTGAGAGCAAAAGTTCAGTTG 942  
 Db 2227139 CCAGTTCCAGCCGCGGACGCAATGCGCCAGCTG 2227172

RESULT 4

US-09-103-840A-1

; Sequence 1, Application US/09103840A  
 ; Patent No. 6294328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEISCHMAN, Robert D.  
 ; APPLICANT: WHITE, Owen R.  
 ; APPLICANT: FRASER, Claire M.  
 ; APPLICANT: VENTER, John C.  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
 ; TITLE OF INVENTION: TUBERCULOSIS  
 ; FILE REFERENCE: 24366-20007.00  
 ; CURRENT APPLICATION NUMBER: US/09/103,840A  
 ; CURRENT FILING DATE: 1998-06-24  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 441529  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; OTHER INFORMATION: H37RV  
 US-09-103-840A-1

Query Match 5.1%; Score 121.2; DB 3; Length 441529;  
 Best Local Similarity 49.5%; Pred. No. 9.8e-24;  
 Matches 433; Conservative 0; Mismatches 423; Indels 18; Gaps 4;  
 Qy 78 GTAACTAAGCGCGCAATCCCTCGATTGCTGATCAACGAGCGGCTCTGTGAGTCTAGCTA 137  
 Db 2229009 GTTGTGCGCCCGGTACAGACCGCTTGGCGCGCGCTCACCGTGTGCGGAAATTCGCGGA 2229068  
 Qy 138 GAGATCTAGATTCCAGCGCGCATCGTTCATACATCGGTGTGTCATATCGGTATCTCAT 197  
 Db 2229069 TGATCGGACTGTCAGATTTTCAGACATTCGCCATAGAGAGGACGTGAGGTGTATGTCG 2229128  
 Qy 198 CGAGAGAGTCACTTCTCTGCTTTTAGCATGCGGAGCGAGCTTGGGTTTCGGAAGAGTC 257  
 Db 2229129 AGACCGGTACGAGAGTTCATCGGACAGCGAGATCTGCCAGCTTCTCGGGAACATGC 2229188  
 Qy 258 CCCAACCAAGCGCTTCGCGGCAATTCCTCAACAAACCTTCGCCGAGCGGACAAATGATA 317  
 Db 2229189 CCGATCCAGCGCGCGCGCTGCGCGGTGAAGCCCTCTGTGCTGGGACAAAGTGG 2229248  
 Qy 318 CGCGCTGCGCGCCCAAGAGACCATCGAC---GCGCGCTTCAGGTTCAGGTCTTGAAGA 374  
 Db 2229249 TCGGTCTGATGCGCGGACGAAAGGCTTTAGCGACACATGCTCTCGAGCCCATCGT 2229308  
 Qy 375 CATCTTTGGACCGCAAGCGCTTAAGAGCGGCTTCGAGCGGCAATCTTAGTTTCCATCAAC 434  
 Db 2229309 CAGATTCACGCGCAGTGAAGCGGCTTTAGCGCGCGCGGCGGAGTAAACCGCTCGGATA 2229368  
 Qy 435 TGTAGGATCCCGCAATGAGGCGGTTCGAATGGCGAGTGGCGCATGTTCCCAAGTTCTA 494

RESULT 3  
 US-09-103-840A-2  
 ; Sequence 2, Application US/09103840A  
 ; Patent No. 6294328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEISCHMAN, Robert D.  
 ; APPLICANT: WHITE, Owen R.  
 ; APPLICANT: FRASER, Claire M.  
 ; APPLICANT: VENTER, John C.  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
 ; TITLE OF INVENTION: TUBERCULOSIS  
 ; FILE REFERENCE: 24366-20007.00  
 ; CURRENT APPLICATION NUMBER: US/09/103,840A  
 ; CURRENT FILING DATE: 1998-06-24  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 4403765  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; FEATURE:  
 ; OTHER INFORMATION: CDC 1551  
 ; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
 ; OTHER INFORMATION: represent a, t, c or g  
 US-09-103-840A-2

Query Match 5.1%; Score 121.2; DB 3; Length 4403765;  
 Best Local Similarity 49.5%; Pred. No. 9.8e-24;  
 Matches 433; Conservative 0; Mismatches 423; Indels 18; Gaps 4;  
 Qy 78 GTAACTAAGCGCGCAATCCCTCGATTGCTGATCAACGAGCGGCTCTGTGAGTCTAGCTA 137  
 Db 2226308 GTTGTGCGCCCGGTACAGACCGCTTGGCGCGCGCTCACCGTGTGCGTAAATTCGCGGA 2226367  
 Qy 138 GAGATCTAGATTCCAGCGCGCATCGTTCATACATCGGTGTGTCATATCGGTATCTCAT 197  
 Db 2226368 TGATCGGACTGTCCAGTTTCCAGCATTCGCAATAGAGAGGACGTGAGGTGTATGTCG 2226427  
 Qy 198 CGAGAGGATCACTTCTGCTGCTTTTAGCATGCGGAGCGCTTGGGTTTCGGAAGAGTC 257  
 Db 2226428 AGACCGGTACGAGCATCCATCGGAGCGGAGATGTCGAGCTTCTCGGGAACATGC 2226487  
 Qy 258 CCGAACCAAGCGCTTCGCGAATTCCTCAACAAACCTTCGCGGAGCGGACATGATA 317  
 Db 2226488 CCGATCCAGCGCGCGCGCTGCGCGGTGAGCGCTCTGTGCTGCGGACAAAGTGG 2226547  
 Qy 318 CGCGCTGCGCGCCCAAGAGCATTCGAC---GCGCGCTTCAGGTTCAGGTCTTGAAGCA 374  
 Db 2226548 TCGGTCTGATGTCGCGGACGAAAGCGCTTACGACCAACATGCTCTGAGCCCATCGT 2226607  
 Qy 375 CATCTTTGGACCGGAGCGTAAAGCGGATCGGACCGCAATCTAGTTTCCGATCAACCA 434  
 Db 2226608 CAGATTCACGCGCTGACGAGCTTTAGCGCGCGCGGAGTGAACCGCTCGGATA 2226667  
 Qy 435 TGTAGGATCCCGCAATGAGGCGGTTCGAATGCGCAAGTGGCGGATGTTCCAGTTCTA 494  
 Db 2226668 GATGCGCTGGAAGATGCGCTGTGCGCACTGATGATGAGTTCACCGCGGGT 2226727  
 Qy 495 CTATTCATATCCCGCAGGATAGCTTTCAGGCTTACGCTTCCTCAAAACATCTCCAC 554  
 Db 2226728 GCACCGGACCGCGGACCGGGTTCCGCTCGGTGTGATCCGCCCATTCGCCACCCCT 2226787  
 Qy 555 CGCGAGCAAGGATATGTCGCTTCTCATCTTCCAGCGGAGCGGTGAGGTGTCAC 614  
 Db 2226788 CCGTAGACGCGCGCGGAATGCTCTGCTCTCGATCCGAGCTCGAGGAGCGTGC 2226847  
 Qy 615 CCCAAGAGCTTACCTCTGTTGAACAGCGGAGGAAACATGTTGATGAGGATCTGCTTGA 674  
 Db 2226848 -----CGAGACCTCGAACACGCGCGGAAACCATGTGCCATGGAATCGGCTTTA 2226898  
 Qy 675 TGCGGATGTTAAAGGGA-----TTTCAGCAAGGCGT--CCAGATAGTGGCGCTTAGATT 728

Thu Mar 18 12:30:59 2004

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5541 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli  
STRAIN: Clinical Isolate EC-625  
US-08-920-812-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;  
Best Local Similarity 50.5%; Pred. No. 2.2e-19;  
Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

1623 CATTGTACCGCCGCTGCTCCAGCCCAAGGTGGCGCTGGATCAACGCTGCTGCTGGGAG 1682  
1687 CATTATCATCTCACTGATCTTATGCGGCGGCTTCTGCGCCCGTTATTAAACAGGTCA 2746  
1683 TTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGTGTAGTTTTCGGGGTTTGG 1742  
2747 TCGTCAATCACTGATCTTATGCGGCGGCTTCTGCGCCCGTTATTAAACAGGTCA 2806  
1743 AATCGGTGCGCTTGGCCCAAAATGTTGATGCGGCTGCTGGGAAATCTCATCGATCGCT 1802  
2807 TATCGGAAGACGCTGCGCAGATTCAGCTGCGCATCGGCGATATGCTGATCAATCTGG 2866  
1803 CCAACTCGGCTGAGAAATCTCAAGTTGTTGATGATCAAGGCTGTTGTCAGCTGCT 1862  
2867 CAGCTCTCTGCTGCTAAATGTCAGATTTTACGCTGCTGCGCTGCTGCTCAAGTTGTC 2926  
1863 CAACTGACGAAGACCAATCAATGCACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1922  
2927 CGCGCGC-TGCGCACCAATCAATACCGAGCTCACGGATCA-----T 2967  
1923 CGCGCAGACCCATGCGAGGCGCATCTGCGAGTGAATGCGCGGTTCTTGGGCGATGT 1982  
2968 CTTTCAGCAACCGAGCTTACGCGCTTTGGCCATTGATGTTCCAGCTGCTGTCACATTT 3027  
1983 CATTGAGCTTGGGACCATATCAATATTGTTTCACTTCAATGCTGCTGCTGCTGCTGCTGCT 2042  
3028 CATTCAATGATGTTAGGCTGTTGAGGTTGGCTTGGTTAGCATTTTTCGGCGCTCAGAC 3087  
2043 TACCTGCTGCTGGC-----CGGGAACCTCTGGAATTCATTCAGATTTTGT 2090  
3088 GAATTTATTCCTTACGATGATCGGTGAATCTTGGGAATGCGGTGAGATTTTTC 3147  
2091 CCGTGACGAGGCTTGGCAGTGTGAGAAAGCAATGAGCCCAAGCAATTTGTTGGCAG 2150  
3148 CGGTGACCAATCCCTGAGCCAGAGGATTAAGCCCAATACAGCCCGGTTATTGCA 3207  
2151 CTGACTGCAACAGTTTCTACCGTCTATCCCGGTTCTTCCACCAAGATTAATGATGG 2210  
3208 GGGTATCCAGCA-----GGCGCTTTTATCCACCGAGGTTTCAAT 3252  
2211 AATAGCTTGGCTGATGATGAGAGCGGCGAGCCCTCTCGCATGAATCAGCCGCT 2270  
3253 TGTACGAAGGTTGATGAATTAACAGCGGAATTTTCCACTCGCGCAGCACTCAACATTT 3312  
2271 CGCTGTGAGCTCTGAGCGGTAGGAAGAAATACCAACGTAAGAGGCTTTTCCAGACGAA 2330  
3313 TTTGCTGCTGCTCTGCGGAGTAAGAGGATCCCGACATTAAGCGCTTACCGCTTTGTA 3372  
2331 CAATGTCAGCAATGCTGATGATGTTTCTTCCAAAGGAGTATCT 2374  
3373 CGCATGAGCCAGCGCAGAGGCGGTTTCTTCCATCGGCTATTT 3416

RESULT 6  
US-08-920-827-20  
; Sequence 20, Application US/08920827  
; Patent No. 5770375

GENERAL INFORMATION:  
APPLICANT: Ohno, Tsuneya  
APPLICANT: Matsunisa, Akio  
APPLICANT: Uehara, Hirotsugu  
APPLICANT: Eda, Soji  
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/920,827  
FILING DATE: 29-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,577  
FILING DATE: 27-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 19036/32420  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5541 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli  
STRAIN: Clinical Isolate EC-625  
JS-08-92-02-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;  
Rest Local Similarity 50.5%; Pred. No. 2.2e-19;  
Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

2y 1623 CATTTGTCAGCGCCGCTCTCAGCCGCCAAGGTGGCGGTGATCAACGTCGTCGTGGCAG 1682  
|||  
db 2687 CATATCCATCTGACGTTGTGCTCTCTGCATTGTGATGGGCTATCTTGACCGACTGAC 2745  
|||  
2y 1683 TTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAGTFTTCGGGGTTTTGG 1742  
|||  
db 2747 TCGTCATATCACTGACTTCTATTCGGGGCGGTTCTCGCGCCGCTATTAAACAGGTCAAT 2805  
|||  
2y 1743 AATCGTGGCGCTTGGCCAAATGTTGATCGGGGTCGTGGGAATCTCATCATCGCT 1802  
|||  
db 2807 TATCGGAAGACGCTTGCACAGATTTCAGCTCGCATCGCGGATGCTGATCAATCTGCG 2866  
|||  
2y 1803 CCAACTCGGGTCAGAACTCCAACTTGTTGATGTAATCAAGGCTGTGTCCAGCTGCT 1862  
|||  
db 2867 CCGCTCTCTCGGTGCTAAATGTAGATTATTCAGCGCTGACGTTCTCTCAAGTGTG 2926  
|||  
2y 1863 CAACTGACGAAGCACCAATCAATGCACTGGTCAACGGTATCCGGCGCGTACTCTCTGCT 1922  
|||  
db 2927 CCGCGGCG-TGGCACCAATCAATACCGACGTCAGCGCATCA-----T 2967  
|||  
2y 1923 CCGCGACGACCCATGCAAGCCCATCTGGCGAAGTGACTGCCCGGTTCTGGGGCGATGT 1982  
|||  
db 2968 CTTTCAGCAACACAGCTTAACGCCATTGGCGGCATTTGATGTTCCAGCTGGTGCATTT 3027  
|||



us-09-105-117k-1.rni

Thu Mar 18 12:30:59 2004

TITLE OF INVENTION: Probe for Diagnosing Infectious Disease

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/362,577C

FILING DATE: 27-MAR-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Hsien

REGISTRATION NUMBER: 33,547

REFERENCE/DOCKET NUMBER: 19036/32420

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

SEQUENCE CHARACTERISTICS:

LENGTH: 5541 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

ORGANISM: Escherichia coli

STRAIN: Clinical isolate EC-625

US-08-362-577C-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;

Best Local Similarity 50.5%; Pred. No. 2.2e-19;

Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

ORIGINAL SOURCE:

ORGANISM: Escherichia coli

STRAIN: Clinical isolate EC-625

US-08-921-177-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;

Best Local Similarity 50.5%; Pred. No. 2.2e-19;

Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

```

QY 1623 CATTGTCAGCGCGCTGCCAGCCCAAGGTTGGCGCTGGATCAACGCTGCTGGCGAG 1682
DB 2687 CATTATCCATCTGAGTTTGTCTCTGCAATTTGTGATGGCTATCTTGACGACTGAC 2746
QY 1683 TTGTGATGACCGCATGCGCATCAAACTGATGTTGATGGTTAGTTTTCGGGGTTTGG 1742
DB 2747 TCGTCATAATCACTGATCTTATGCGGGCGCGTCTCGCGCCGGTTTAAACAGTCAAT 2806
QY 1743 AATCGGTGGCTTCGCCCAATGTTGATCCGGGCTGTGGGAAATCTCATCGATCGCT 1802
DB 2807 TATCGGAAGACCGCTGCCACAGATTCAGCTGCCATCGCGGATATGCTGATCAATCTGG 2866
QY 1803 CCAACTCGGGGTGAGAAATCTCAAGTTGTGATGTAATCAAGCTGTTGTCAGCTGCT 1862
DB 2867 CCAGCTCTCGGTCTAAATGTCAGATTATTCAGCGCTGCGAGTTCTCTCAAGTTGTC 2926
QY 1863 CCACTGACGAAGCAACATCAATGATGTTGATGTAATCAAGCTGTTGTCAGCTGCT 1922
DB 2927 CGCGCGGC-TGGCACCATAATACCGAAGTCACGCGATCA-----T 2967
QY 1923 CGCGCAGCCCATGCAAGCGCATCTCGCAAGTGAAGTCTGCGCGCTTCTCGGGGATGT 1982
DB 2968 CTTTCAGAACAGCTTAACGCCAATTTGGCCCAATTTGTCACCGCTGCTGCAATTT 3027
QY 1983 CATTGAGCTTGGGACCATATCATATATGTTCACTTCAATCGCCCTCAGACAGGACT 2042
DB 3028 CATTCAATAAGTTGAGCTGTTGAGGTTGGCTTCGGTAAGCAATTTTCGGCGTCACAC 3087
QY 2043 TACCTTGGCTGGC-----GGGGACCTCTGGAATTCATCGAGATATTTGT 2090
DB 3088 GAACCTTATTCCTTCAGATGATCGGTGAATCTTGGGAATGCGGTGAGATATTTTC 3147
QY 2091 CCGTGACGAGCGCTCGCAAGTGTGAGAAAGCAATGACGCCAAGCAATTTGTTGGCGAG 2150
DB 3148 CGGTGACGATCTTCCTGAGCCAGGAGTAAGCAATACAGCCCAAGCGCTTATTTGCA 3207
QY 2151 CTGACTGCAACAAAGTTCTCACCGTCAATCGCCCGTTCTTCCACCCAGATTAATGATGG 2210
DB 3208 GGGTATCCAGCA-----GGCGGCTTTTATCCACCCAGCGGTTCAGTAAT 3252
QY 2211 AATAGCTTGGCTGATGATCAGAAAGCGGCGAGCGCTCTCGCGCATGAATCAGCGCGCT 2270
DB 3253 TGTACGAAGTTGATGAATTAACAGCGGAATTTTCACTCGCGGAGCACTCAACCATTT 3312
QY 2271 CCGGTGTGAGCTCTGGAAGCTGAGGAAGAAATACCCAGTAAGAGCGCTTTCAGAGCGAA 2330
DB 3313 TTTCGCTTCGCTCTCGCGAGTAAGAGAGATCCCGACATAAAGCGCTTACCGTTTGTGA 2374
QY 2331 CAATGTACGCAATCGGTACATGTTTCTTCCAAAGGATATCT 2374
DB 3373 CGCATGAGCCAGCGAGGCGGTTTCTTCCATCGCGGTAATTT 3416

```

RESULT 8

US-08-362-577C-20

; Sequence 20, Application US/08362577C

; Patent No. 5807673

; GENERAL INFORMATION:

; APPLICANT: Onno, Tsuneya

; APPLICANT: Matsuhisa, Akio

; APPLICANT: Uehara, Hirotsugu

; APPLICANT: Eda, Soji

Best Local Similarity 50.5%; Pred. No. 2.2e-19;  
Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;  
1623 CATTGTGACGCGCCCTGCGCAAGTGTGGCGCTGGATCAACGCTGTGTGGCAG 1682  
1687 CATTATCCATCTGACGTTGTGCTCTCTGCAATTTGTGATGGGCTATCTTGACGACTGAC 2746  
1683 TTGTGATGACGCGATTCGCCATCAAACTGATTTGTATGGGTTAGTTTCGGGGTTTTGG 1742  
2747 TCGTCAATATCACTGATCTTATGCGGGCGGTTCTCGCGCCGTTATTAAACAGGTCATT 2806  
1743 AATCGGTGCGCTTGGCCCAATGTTGATGCGGGCGGCTGTGGGAATCTCATCGATCGCT 1802  
2807 TATCGGAGACGCTCTGCGCAGATTCAGTCTGCCATCGCGATATGCTGATCAATCTGGC 2866  
1803 CCACTCGCGGTCGAGAAATCCCAAGTTGTGAGTGAATCAAGGCTGTGTCCAGTGTCT 1862  
2867 CCACTCTCTCGGTGTTAATGTCAGATTAATTCAGGCGCTGCAGTTCTCTCAAGTTGTC 2926  
1863 CAATGACGAGACCAATCAATGTCATGCTGTCAGGTTATCGCGCGCTGACTCTCTCTTGTCT 1922  
2927 CGCGGCGC-TGSCACCAATCAATACCGAGCTCACCGCATCA-----T 2967  
1923 CGCGGAGCACCATGCAAGCGCATCTGCGGCAAGTGTGCTGCGGCTTCTTGGCGGATGT 1982  
2968 CTTTTCAGCAACCGAGCTTAACGCCATTTGCGGCATTTGATTTGTCAGGCTGTGTCCTATTT 3027  
1983 CATTGAGCTTGGGACCATATCAATATTTTCAAGTTCAACATGCTTCCATGCGCTCAGACGGA 2042  
3028 CATTCAATAGTGTAGGCTGTGAGTTGGCTTGGTTAAGCATTTTTCGGCGCTCAGACCCAC 3087  
2043 TACCTTGGCTGGC-----GCGGGAACCTCTGGAATTCATCGAGATTTTGT 2090  
3088 GAACTTTATTTCCCTTACAGATGATCGGTAATCTTTCGGAATGCGGTTGAGATTTTTC 3147  
2091 CGGTGACGAGGCTTCCGCAAGTGTGAGAAAGCAATGAGCGCAAGCAATTTGTGCGAG 2150  
3148 CGGTGACCAATCCCTGAGCGGAGAGTAAGCAATACAGCCAGCGGTTATTTGCA 3207  
2151 CTGACTGCAACAAAGTTCTCACCGCTCATCGCGGTTCTCCACCAACGATTAATGATGG 2210  
3208 GGGTATCCAGCA-----GGCGGCTTTTATCCACCGAGCGGTTCAAGTAAT 3252  
2211 AATAGCTTGGCTGATGATCAAGAGCGGCGAGCGCTCTCTCGGCATGAACTCAGCGGCT 2270  
3253 TGTACGAAGGTTGATGAATTAACAGCGGAATTTTCACTCGCGGCAACTCAACCAATTT 3312  
2271 CGGCTGTGAGCTCTGAGCGCTGAGCAAGAAATACCAAGAGGCTTTCCAGAGCGCA 2330  
3313 TTTGGCTCGCTCTGGCGAGTGAAGAGGAGATCCCGACATAAAGCGGCTTACCGCTTTGTA 3372  
2331 CAATGTCAGCAATGCTGATGATGTTTCTTCCAAAGGAGTATCT 2374  
3373 CGCATGAGCCGAGCGAGCGGCTTTCTTCCATCGCGGTTATTT 3416

Query Match 4.2%; Score 100.8; DB 2; Length 5541;  
US-08-920-828-20  
; Sequence 20, Application US/08920828  
; Patent No. 5853998  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotsugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,828  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,577  
; FILING DATE: 27-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 19036/32420  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5541 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli  
; STRAIN: Clinical Isolate EC-625  
US-08-920-828-20

RESULT 10  
US-09-489-039A-5370/c  
; Sequence 5370, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 5370  
; LENGTH: 1095  
; TYPE: DNA

REF: DIA

LENGTH

TYPE: DNA

```
; SEQ ID NO 791
:      LENGTH: 936
```

```
; SEQ ID NO 791
:      LENGTH: 936
```

ORGANISM: Proteus mirabilis  
US-09-543-681A-791

Query Match 3.1%; Score 73.4; DB 4; Length 936;  
Best Local Similarity 44.9%; Pred. No. 1.2e-11;  
Matches 324; Conservative 0; Mismatches 391; Indels 6; Gaps 1;  
206 ATCACTCTCTCTCTTTTATGATGGGAGCAGCTTTGGTTTGGGAGAGATCCCAACA 265  
Db ATCAATTCACCACTTTTATGCTCATCGGCAATTTGTAGATGAGGATCATACAATGTT 741  
266 AGCCTCGCGGATGCTTCCCAAACTTCGCGGAGGAGGAGATGATGCGGCTG 325  
Db GAGCCTTTGTTAGCTTAATGAACAAGGCTTCGAAGATTAACATATGCGAGGACCA 681  
326 CGCCCAACAGGACCATCGACGCGCGCTCCAGGTCACGGTCTTTGAAGACATCTTTGGGA 385  
Db CTACCGGAGATACCGAGTTTTCCTGTGAAGAGGCTGATGATATCATCAAGATGG 621  
386 CGAAGCGTAAGACGGG-----CATGCGACCAATCTAGTTTCCCATCAACATGATG 439  
Db TCAATGCGACAGCGGGGCTTTTAAACAGAGATGATTTAGTACCCCAATTTGCAAAATAG 561  
440 GCATCCGCGCAATGAGGGGTTGCAATGCGCAAGTGCGCATGTTCCAAAGTTCTACTACT 499  
Db CGTTGGCGAAATACAGGGGATGCCCAAAAGATGATCTATGCGGCTTAATGTCGACT 501  
500 TCATCTCCGCGACAGGATGATCTTCACGGTTTACCGCTCTTAAACATCTCCACGCGCG 559  
Db AGACAGTAGGCAATGCTTGGGCTGAATCTGATAGCAACCAACCTTACCACTCTT 441  
560 AGCAAGATATGTTGGGCTTCACTTCCAAAGCGAGGTGAGGTTGCTTCAACCCCA 619  
Db AATCGCTCTGCTGACGAGTTTCACTCTTGAATATTAAGCGGATAGGAGTTGT 381  
620 GAAGCTACTCTGTCACACGAGGAGAACCATGATGATAGCGAATCTGCGTTGATGGGG 679  
Db GTTAATACCGGTCGAAGCAGGCAATACCAAGTAGTAACTATCGCGCTTGAAGCT 321  
680 ATGTTTAAAGGATTTTACGACGCGTCCAGATAGTTGCGCTTTAGTTTCTGCTTGACG 739  
Db AAAGAGATTAAGAGGCGTCGAGCCACTATTTTCATCACTAAACCACTGCTCTTCTAAC 261  
740 AACACATTTTCGCGCTGCTTGAACAAGCACTTCAACCGCTTCGGTTGCTTTGGCGGT 799  
Db ATTTGCACTGATGATGATGATGAGTAGCTTTGCGGCTGCTCTGCGGTTGGGCGGA 201  
800 TGGTGGCGGATACCAACATCTGACCCACGCTGATGCTGAGAGCTTTAAAGCGCTGACTC 859  
Db ACCGTAGTACTTAATAAGGTTTGCCTAATAGTTTCTCAAGTTGTTTATTTCTTTGTGA 141  
860 ACCGCGAGGGGAAATGAAAGGCTTAAGGAGGCGGCTTGAAGCTGCTTCAATGATG 919  
Db ACGGAGATTTGCTTAATACATAGTTTGTGGGCTGCTTCAAAACCAAGCTCTCGAATA 81  
920 A 920  
80 A 80

RESULT 13  
US-09-252-991A-7644/c  
Sequence 7644, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 7644  
LENGTH: 900  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7644  
Query Match 2.7%; Score 65.2; DB 4; Length 900;  
Best Local Similarity 46.3%; Pred. No. 3.2e-09;  
Matches 412; Conservative 0; Mismatches 453; Indels 25; Gaps 5;  
87 GCAGCAATCCCTCGATTTGCTCATCAACGACGCGGCGGAGCAGTCTCGTGAGCGACGAGCAGTTGCG 827  
Db GCACCAAGCGTTCGCGGCGGCGGCGGAGCAGTCTCGTGAGCGACGAGCAGTTGCG 827  
147 ATTCCAGGCGCATCGTTGCGCAATATCATCGTGTGTCTCATGAGGATCTCTCATCGAGAG-- 204  
Db CGCGTTGCGCGAGTAGTGCAGTACACGCTACGCTGATGACCTGCGCGGCGAGCAGCT 767  
205 -GATCATTCTCTCTCTTTTATGATGGAGCAGCTTGGTTTCGGGAGAGAAAGTCCCCAAC 263  
Db CGACCAAGTTCACCGCGCGCGCTTTCGACCTGCGCTTCGCGCACCAGGCGCCGAGC 707  
264 CAAGGCTCGCGCAATTCCTCTCAACAAACCTTCGCGCGACGCGCAATGATAGCGGC 323  
Db GAGACCGCGCGGCTGAGGCGGACGAACTTCGAGGAGCGGCGAGAGTGTGATGA 647  
324 TCGCGCGCGACGAGCAATATGAGCGCGCGCTTCAGTCTTGAAGCAGCATCTTTGG 383  
Db AACCGCGCTCGA-----CGCGAGGCTCTTGAAGAGCGGTGTTCGACGAGTGTGTCG 593  
384 GACCGAGCGTGAAGCGGCGCATCGCAGCCCAATCTAGTTTCCCATCAACATGATAGGAT 443  
Db GCGCGAAACAGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGA CCGCGGGGAAAGT 533  
444 CCGCG-----AATGAGGGGGTTCGAATGCGCAAGTGGCGCATGGTTCCAGTTCTACTA 497  
Db CCGCGCAATGAATCCCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 473  
498 CTTCAATCCCGCGCGCGGATTTAGCTTTCAGGGTTTACCGCTTCTTAAACATCTCCACGCC 557  
Db AACG 413  
558 GCAGCAGGATATATGTTGCGCTTTCATCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCG 617  
Db GCATGCGCTTGAAGCG 353  
618 AAGAAGCTTACTCTGTTGAACACGCGGAGGAAACCATGTGATAGGGAATCTGCGTTGA--TG 676  
Db CGCGCGCAATTCGCCACCG 293  
677 GCGATGTTTAAAGGATTTTACGAGCGGCTTCAGATGTTGCG-----CTTAGTT 727  
Db GCGATGCG 233  
728 TCTGCTTGCAGCAACACCATTTTTCGCGCTGCTTTCACAGGACTTTCACCGCTTTCGCTT 787  
Db TCGCTTCCAGCAGCG 173  
788 GCTTTGCGCGGTTGGGTGCGGATACCAACATCTGACCGCGCGCGCGCGCGCGCGCGCG 847  
Db GGTGCGCGCGGCTCTCTCG 113  
848 AGCGCTGATCTACCGCGCGCGGAGGAAATGGAAGGCTTAAGGAGCGCTTCCAGCTG 907  
Db ATCCGCTCGGATACCG 53  
908 CTTTCAATCAATGATGAGAGCAAGTGTCCAGTTGATGGGCTTCATGA 957  
Db CCGTCTCCACACTGCG 3

Thu Mar 18 12:30:59 2004

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 PRIOR FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 7908  
 LENGTH: 699  
 TYPE: DNA  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-7908

Query Match 2.6%; Score 51.8; DB 4; Length 699;  
 Best Local Similarity 54.6%; Pred. No. 2.9e-08;  
 Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;  
 QY 1455 TGAATGCAATCGTGTGACCTGTTGAACCCGAATGCGTATTTCGACGGGTTGTTGA 1514  
 Db 434 TGGCGGCACTGGCGGTACCCCTGCTCAACCCCGACGCTCTATCTCGATACCGTACTCTGA 493  
 QY 1515 TCGGCGGCGTGGCGGCGCAATACGGCGACACCGGACGGTGGATTTCGCGCTGGCGGT 1574  
 Db 494 TCGGTTGCTCGGCGGCGGCGGAGG-----CCGCGCGGCGGCGGCTATGCGCTTCGCGGCGG 547  
 QY 1575 TCGCGGCAAGCTGATCTGTTCCGCTGGTGGTTCGCGCGACGACGATTTGTCACGCC 1634  
 Db 548 CCAGCGCTCTCGTGTGATGTTTCGCGCTTCGCGCTTCGCGCGGCGATGGCTGGCGGCT 607  
 QY 1635 CGCTGTCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1694  
 Db 608 GGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1725  
 QY 1695 CATGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1725  
 Db 668 GCATGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 698

Search completed: March 16, 2004, 03:00:33  
 Job time : 198.837 secs

RESULT 14  
 US-09-724-623-24/c  
 Sequence 24, Application US/09724623  
 Patent No. 6476209  
 GENERAL INFORMATION:  
 APPLICANT: Glenn, Matthew  
 APPLICANT: Lubbers, Mark W  
 APPLICANT: Dekker, James  
 TITLE OF INVENTION: Polynucleotides, materials incorporating  
 TITLE OF INVENTION: them, and methods for using them.  
 FILE REFERENCE: 1048U1  
 CURRENT APPLICATION NUMBER: US/09/724,623  
 CURRENT FILING DATE: 2000-11-28  
 NUMBER OF SEQ ID NOS: 124  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 24  
 LENGTH: 1002  
 TYPE: DNA  
 ORGANISM: Lactobacillus rhamnosus  
 US-09-724-623-24

Query Match 2.7%; Score 64; DB 4; Length 1002;  
 Best Local Similarity 47.9%; Pred. No. 7.8e-09;  
 Matches 279; Conservative 0; Mismatches 270; Indels 33; Gaps 2;  
 QY 1785 AAATCTCATCGATCGGCTCCCACTCGCGCTCAGAAACTCCCAAGTTGTTGAGTGAATCAA 1844  
 Db 994 AAATATCATCAATTTGTTGATCTCTTCAGCAGTAAGGTCAAGTTCGGTTGCTTTAA 935  
 QY 1845 GCGCTGTTGTCAGCTCTCACTGACGAAGCAACCAATCAATGCTGCTGCTGCTGCTGCTGCT 1904  
 Db 934 GGTATCTGAAGGTGTTCACTGAGGTGCTGCCAATGATCACTGTTGACAA-----881  
 QY 1905 CGCGGTACTCTCTGCTGCGCAGCACCCCATGACAGGCCATCTGCGAAGTGACTGCC 1964  
 Db 880 -----CGGATCCCGTAACACACCGCCAGGCCATTGACTCAGGGTTTGT 833  
 QY 1965 CGGTTCTGCGGCGATGATGATGAGCTTGGGACCATATCAATATTTGTCAGGTTCAACA 2024  
 Db 832 CAGATCATGCGCAATTTCTTAAGCGCATTTAGTTGTTTAAACACAGCCCTCTTTGCCCT 773  
 QY 2025 TGCCCTCAGACAGGAGCTTACCGTGGCGGCGGAGCCCTCTGGAATTCATCGAGAT 2084  
 Db 772 TAGCAAAAGTGGCTTGTGTTGATGATTTTGAAGTATCGGAATTCCTTTAGGT 713  
 QY 2085 ATTGTCCGTGAGCAGCGCTGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCATTTGT 2144  
 Db 712 AGCGATCTGATAAAGCGCTTCTGATAAGGTTCGTATGATCAACCGCGCACCATCAG 653  
 QY 2145 TGGCAGCTGACTGCAACAAAGTTCTCACCGTCATCGCGGTTCCCTCCACCCACGATTAA 2204  
 Db 652 CTTTTATGATCGATCAA-----GCCGAGCGTTTCAGCGGTGGGATTAA 608  
 QY 2205 TGATGGAATAGCTTGGCTGATGATCAGAAAGCGGAGCCCTCTCCCGCATGAATCAG 2264  
 Db 607 ACATGTTGTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 548  
 QY 2265 CGGCTCCGCTGAGCTCTGAGCGGTAGGAGAAATACCAAGTAAAGAGAGCCCTTTCCAG 2324  
 Db 547 TTGCTTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 488  
 QY 2325 AGGCAACAATGTCAGCAATGCGTACATGTTTCTTCCAAAG 2366  
 Db 487 CAGCAGCGTTTATCCAGCGCATTAACCGTCTCTTCCAAAG 445

RESULT 15  
 US-09-252-991A-7908  
 Sequence 7908, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.

Result No.	Score	Query		Length	DB	ID	Description
		Match	Length				
1	2374	100.0	2374	14	US-09-198-232-24		Sequence 24, Appl
C	2	2374	100.0	3309400	9	US-09-738-626-1	Sequence 1, Appl
	3	993	41.8	993	10	US-09-746-660A-47	Sequence 47, Appl
C	4	870	36.6	870	9	US-09-738-626-3456	Sequence 3456, Ap
5	822	34.6	822	10	US-09-746-660A-51	Sequence 51, Appl	
6	711	29.9	711	14	US-10-166-142-7	Sequence 7, Appl	
7	708	29.8	708	9	US-09-738-626-3455	Sequence 3455, Ap	
8	636.8	23.4	712	14	US-10-166-142-9	Sequence 9, Appl	
C	9	627	26.4	627	9	US-09-738-626-3454	Sequence 3454, Ap
	10	485.2	20.4	993	12	US-10-627-476-669	Sequence 669, App
C	11	305.8	12.9	1095	12	US-10-282-122A-17521	Sequence 17521, A
C	12	302.8	12.8	879	12	US-10-282-122A-17523	Sequence 17523, A
C	13	164.4	6.9	1026	15	US-10-369-493-35374	Sequence 35374, A
C	14	164.4	6.9	1026	15	US-10-369-493-38826	Sequence 38826, A
C	15	164.4	6.9	1028	15	US-10-369-493-38124	Sequence 38124, A

QY	121	GTCTGTGAGTCTAGTAGAGATCTAGATTCACGCGCCATCGTTGCGCAATACATCGGTGT	180
Db	121	GTCTGTGAGTCTAGTAGAGATCTAGATTCACGCGCCATCGTTGCGCAATACATCGGTGT	180
QY	181	GTCAATGGGTATCTCATCGAGAGATCATCTTCTCTGCTTTTAGCATGGGAGCTTG	240
Db	181	GTCAATGGGTATCTCATCGAGAGATCATCTTCTCTGCTTTTAGCATGGGAGCTTG	240
QY	241	GGTTTCGGAGAGTCCCAACCAAGGCTCGGCGAATTTGCTCCACAAAACCTTCCGC	300
Db	241	GGTTTCGGAGAGTCCCAACCAAGGCTCGGCGAATTTGCTCCACAAAACCTTCCGC	300
QY	301	CGACGGACAATGATAGCGCTCGCGCCACACAGAACATCGACGCGCTCGCAGGTC	360
Db	301	CGACGGACAATGATAGCGCTCGCGCCACACAGAACATCGACGCGCTCGCAGGTC	360
QY	361	ACGGTCTTGAAGCAATCTTTGGGACCGAAGGTAAAGACGGGATCGAGCCCAATCTAG	420
Db	361	ACGGTCTTGAAGCAATCTTTGGGACCGAAGGTAAAGACGGGATCGAGCCCAATCTAG	420
QY	421	TTTCCCATCAACCATGAGGATCCGCAATGAGGGGTTGCAATGGCGCAAGTGGCGAT	480
Db	421	TTTCCCATCAACCATGAGGATCCGCAATGAGGGGTTGCAATGGCGCAAGTGGCGAT	480
QY	481	GGTTCCAAAGTCTTACTACTTCAATCCCGCCACGGGATAGCTTACGGGTTACCGTTC	540
Db	481	GGTTCCAAAGTCTTACTACTTCAATCCCGCCACGGGATAGCTTACGGGTTACCGTTC	540
QY	541	TAAACATCTCACGCGCAGCAGGATAGTGTGGCTTCATCTTCCAGCGCAGCGT	600
Db	541	TAAACATCTCACGCGCAGCAGGATAGTGTGGCTTCATCTTCCAGCGCAGCGT	600
QY	601	GAGGTTGTCTCACCCCAAGAGCTACTCTGTGAAACACGGAGGAACCAATGTGATAG	660
Db	601	GAGGTTGTCTCACCCCAAGAGCTACTCTGTGAAACACGGAGGAACCAATGTGATAG	660
QY	661	CGAATCTCGCTGATGGCGATGTAAACGGGATTTACGAGGGTCCAGATAGTTGGC	720
Db	661	CGAATCTCGCTGATGGCGATGTAAACGGGATTTACGAGGGTCCAGATAGTTGGC	720
QY	721	TTTAGTTCTGTGAGCAGCACATTTTCGCGCTGTGACACAGGACTTCAACCGC	780
Db	721	TTTAGTTCTGTGAGCAGCACATTTTCGCGCTGTGACACAGGACTTCAACCGC	780
QY	781	TTTCGGTTCCTTGGCGGTTGGTGGCGATACCAACACTCGACCCACGTGATGTCGAG	840
Db	781	TTTCGGTTCCTTGGCGGTTGGTGGCGATACCAACACTCGACCCACGTGATGTCGAG	840
QY	841	AGCTTTAAGCGCTGACTCACCGCGAGGGGAAATGGAAGGCTAAGAGGCGCCTTC	900
Db	841	AGCTTTAAGCGCTGACTCACCGCGAGGGGAAATGGAAGGCTAAGAGGCGCCTTC	900
QY	901	GRAGCTGCTTCAATCATGATGAGCAAAAGTGTCCAGTTGAATGGGGTTCAATGAAGT	960
Db	901	GRAGCTGCTTCAATCATGATGAGCAAAAGTGTCCAGTTGAATGGGGTTCAATGAAGT	960
QY	961	ATATTAAACCAATGTAAGAACCAATCAATTTTACTTAAGTATCTCCATAGTTCAGATGT	1020
Db	961	ATATTAAACCAATGTAAGAACCAATCAATTTTACTTAAGTATCTCCATAGTTCAGATGT	1020
QY	1021	GATCATGGAAATCTTCAATACAGTCTGCTTTGGGGCCAGTCTTTTACTGTCCATCGG	1080
Db	1021	GATCATGGAAATCTTCAATACAGTCTGCTTTGGGGCCAGTCTTTTACTGTCCATCGG	1080
QY	1081	ACCGAGAAATGACTGTGATTAACAGAAATTAAGCGGAGGACTTCAATCGGTTCT	1140
Db	1081	ACCGAGAAATGACTGTGATTAACAGAAATTAAGCGGAGGACTTCAATCGGTTCT	1140
QY	1141	TCTGTGTGTTTAATTTCTGACGCTTTTGTTCATCGCGGACCTTGGCGGTTGATCT	1200
Db	1141	TCTGTGTGTTTAATTTCTGACGCTTTTGTTCATCGCGGACCTTGGCGGTTGATCT	1200
QY	1201	TTTGTCCAAATCGCGCGCATCGTGTCTCGATTAATTAATGCTGGGTGGCATCGCTTACCT	1260
Db	1201	TTTGTCCAAATCGCGCGCATCGTGTCTCGATTAATTAATGCTGGGTGGCATCGCTTACCT	1260
QY	1261	GTATGTTTGGCGGTCAATGGGAGCAAGACCGCATGACAAACAAAGGTGGAAGCGCAC	1320
Db	1261	GTATGTTTGGCGGTCAATGGGAGCAAGACCGCATGACAAACAAAGGTGGAAGCGCAC	1320
QY	1321	GATCATTTGAAGAAACAGAAACCAACCGTCCCGATGACACGCTTTTGGCGGTTCGGCGGT	1380
Db	1321	GATCATTTGAAGAAACAGAAACCAACCGTCCCGATGACACGCTTTTGGCGGTTCGGCGGT	1380
QY	1381	GGCCACTGACACGCGCAACCGGTGCGGTTGAGGTGAGGTGAGTGAATGAGCGGTTTG	1440
Db	1381	GGCCACTGACACGCGCAACCGGTGCGGTTGAGGTGAGGTGAGTGAATGAGCGGTTTG	1440
QY	1441	GATTAAGCCCATGTTGATGAGCAATCGTGTGACCTGTTGAAACCCGAATGCTATTGGA	1500
Db	1441	GATTAAGCCCATGTTGATGAGCAATCGTGTGACCTGTTGAAACCCGAATGCTATTGGA	1500
QY	1501	CGCGTTTGTGTTTATCGCGCGCGTTCGCGCGCAATACGGCGACACCGGACGGTGGATTT	1560
Db	1501	CGCGTTTGTGTTTATCGCGCGCGTTCGCGCGCAATACGGCGACACCGGACGGTGGATTT	1560
QY	1561	CGCGCTGCGCGGTTTCGCGCGCAAGCTGATCTGTGTTCCCGTGGTTCGCGCGGAGC	1620
Db	1561	CGCGCTGCGCGGTTTCGCGCGCAAGCTGATCTGTGTTCCCGTGGTTCGCGCGGAGC	1620
QY	1621	AGCATTTGACGCGCGTTCGCGCGCAAGGTTGCGGCTGGATCAAGCTGCTGCTGGC	1680
Db	1621	AGCATTTGACGCGCGTTCGCGCGCAAGGTTGCGGCTGGATCAAGCTGCTGCTGGC	1680
QY	1681	AGTTGCTGATCAGCCATTTGGCCATCAAACTGATGTTGATGGGTTAGTTTCGCGGGTTT	1740
Db	1681	AGTTGCTGATCAGCCATTTGGCCATCAAACTGATGTTGATGGGTTAGTTTCGCGGGTTT	1740
QY	1741	GGATTCGGTGGCTTCGCGCGCAATGTTGATGCGCGGCTGTTGGAAATCTCATCGATCGC	1800
Db	1741	GGATTCGGTGGCTTCGCGCGCAATGTTGATGCGCGGCTGTTGGAAATCTCATCGATCGC	1800
QY	1801	CTTCAACTCGCGGCTCAGAAACTCCAAAGTTTGTAGTGAATCAAGGCTGTTGTCAGCTG	1860
Db	1801	CTTCAACTCGCGGCTCAGAAACTCCAAAGTTTGTAGTGAATCAAGGCTGTTGTCAGCTG	1860
QY	1861	CTCAACTGACGAGCAACCAATCAATGCACTGTTCAAGGCTGTTGTCAGCTG	1920
Db	1861	CTCAACTGACGAGCAACCAATCAATGCACTGTTCAAGGCTGTTGTCAGCTG	1920
QY	1921	CTTCAACTGACGAGCAACCAATCAATGCACTGTTCAAGGCTGTTGTCAGCTG	1980
Db	1921	CTTCAACTGACGAGCAACCAATCAATGCACTGTTCAAGGCTGTTGTCAGCTG	1980
QY	1981	GTCAATGAGCTTGGGACCAATCAATATTTGTTCAAGTTCAACATGCTCCAGAGGAA	2040
Db	1981	GTCAATGAGCTTGGGACCAATCAATATTTGTTCAAGTTCAACATGCTCCAGAGGAA	2040
QY	2041	CTTACCTGCTGCGCGGAAACCTCTCTGGAATTTCCATCGAGATTTTGTTCGTGAGAG	2100
Db	2041	CTTACCTGCTGCGCGGAAACCTCTCTGGAATTTCCATCGAGATTTTGTTCGTGAGAG	2100
QY	2101	GCCTCGGCAAGTGTGAGAAAGCAATGACCGCAAGCAATTTGTTGGCAGCTGACTGCA	2160
Db	2101	GCCTCGGCAAGTGTGAGAAAGCAATGACCGCAAGCAATTTGTTGGCAGCTGACTGCA	2160
QY	2161	CAAGTTCTCACCGTCACTCGCGCGGTTCTCCACCAACCAACCAACCAACCAACCA	2220
Db	2161	CAAGTTCTCACCGTCACTCGCGCGGTTCTCCACCAACCAACCAACCAACCAACCA	2220
QY	2221	CTGATGAATCAGAGCGGAGCCCTCTCTCGCCATGAACCTGAGCGGCTCGCTGTGAG	2280
Db	2221	CTGATGAATCAGAGCGGAGCCCTCTCTCGCCATGAACCTGAGCGGCTCGCTGTGAG	2280
QY	2281	CTCTGGACCGTAGGAGAAATACCCACGTAAAGAGGCTTTTCCAGACGCAACAAATGTCAG	2340
Db	2281	CTCTGGACCGTAGGAGAAATACCCACGTAAAGAGGCTTTTCCAGACGCAACAAATGTCAG	2340



Db 2281 CTCTGACCGTAGGAAGAAATACCAACGTAAGAGCCCTTTCCAGACCAACATGTCAGC 2340  
QY 2341 CAATGGTACATGTTCTTCCAAAGGAGTATCT 2374  
Db 2341 CAATGGTACATGTTCTTCCAAAGGAGTATCT 2374  
RESULT 2  
US-738-626-1/c  
; Sequence 1, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIALI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09738, 626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 1  
; LENGTH: 3309400  
; TYPE: DNA  
; ORGANISM: *Corynebacterium glutamicum*  
US-09-738-626-1  
Query Match 100.0%; Score 2374; DB 9; Length 3309400;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCATTTGCTGAAGGTGTTACTCTGCTGCGCCCAATTCCTGCGGCGAAGAGTGAAGAAC 60  
Db 132968 CCAATTTGCTGAAGGTGTTACTCTGCTGCGCCCAATTCCTGCGGCGAAGAGTGAAGAAC 132969  
QY 61 CCGAACCTTTTCAGAGTAAGTAAAGCGCGCAATCCCTCGATTGCTGCATCAAGAGCGGC 120  
Db 132908 CCGAACCTTTTCAGAGTAAGTAAAGCGCGCAATCCCTCGATTGCTGCATCAAGAGCGGC 132949  
QY 121 GTCGTGAGTCTAGCTAGAGATCTAGATTCAGAGCGCCATCGTTGCCAATACATCGGTG 180  
Db 132948 GTCGTGAGTCTAGCTAGAGATCTAGATTCAGAGCGCCATCGTTGCCAATACATCGGTG 1329789  
QY 181 GTCAATGGGTATCTCATGAGAGAGTCACTTCTCTGCTTTAGCATGGGAGCAGTTG 240  
Db 1329788 GTCAATGGGTATCTCATGAGAGAGTCACTTCTCTGCTTTAGCATGGGAGCAGTTG 1329729  
QY 241 GGTTCGGGAAGAGTCCCAACCAAGGCTCGGGCAATTCGCTCACCRAAACCTTCGCG 300  
Db 1329728 GGTTCGGGAAGAGTCCCAACCAAGGCTCGGGCAATTCGCTCACCRAAACCTTCGCG 1329669  
QY 301 CGACGGGACATGATATACCGCTCGCGCCCAACAGGACCATCGAGCGCGCCGTCAGATC 360  
Db 1329668 CGACGGGACATGATATACCGCTCGCGCCCAACAGGACCATCGAGCGCGCCGTCAGATC 1329609  
QY 361 ACGGTCTTGAAGCACATCTTTGGGACCGAAGCGTAAGAGCGGCGATCGAGCCCATCTAG 420  
Db 1329608 ACGGTCTTGAAGCACATCTTTGGGACCGAAGCGTAAGAGCGGCGATCGAGCCCATCTAG 1329549

QY 421 TTTCCCATCAACCATGTAGCATCCCGCAATGAGGGGTTGCAATGGCCAAAGTGGCGCAT 480  
Db 1329548 TTTCCCATCAACCATGTAGCATCCCGCAATGAGGGGTTGCAATGGCCAAAGTGGCGCAT 1329489  
QY 481 GGTTCCAAAGTCTTACTTCAATCCCGCAACGGGATTTAGCTTCAAGGGTTACCGCTCC 540  
Db 1329488 GGTTCCAAAGTCTTACTTCAATCCCGCAACGGGATTTAGCTTCAAGGGTTACCGCTCC 1329429  
QY 541 TAAACATCTCCAGCGCGCAGCAAGGATAATGTGTGGCTTTCATCTTCCAAAGCGCAGCGT 600  
Db 1329428 TAAACATCTCCAGCGCGCAGCAAGGATAATGTGTGGCTTTCATCTTCCAAAGCGCAGCGT 1329369  
QY 601 GAGGTTGTCTCCACCCCAAGAAGCTACCTCGTTGAACAACGGGAGGAAACCATGTGGATAG 660  
Db 1329368 GAGGTTGTCTCCACCCCAAGAAGCTACCTCGTTGAACAACGGGAGGAAACCATGTGGATAG 1329309  
QY 661 CGAATCTGCTGTGATGGCGATGGTTAAACGGGATTTACAGCAAGCGCTCCAGATAGTTCGCG 720  
Db 1329308 CGAATCTGCTGTGATGGCGATGGTTAAACGGGATTTACAGCAAGCGCTCCAGATAGTTCGCG 1329249  
QY 721 TTTAGTTCTGCTTCCAGCAACACCAATTTTCCGCGCTGCTTGCACAAGGACCTTCACCGCG 780  
Db 1329248 TTTAGTTCTGCTTCCAGCAACACCAATTTTCCGCGCTGCTTGCACAAGGACCTTCACCGCG 1329189  
QY 781 TTCGTTGCTTTGGCGGTTGGGTGCGGATACCAACAACCTCGACCCACCGTGAATGCTCGAG 840  
Db 1329188 TTCGTTGCTTTGGCGGTTGGGTGCGGATACCAACAACCTCGACCCACCGTGAATGCTCGAG 1329129  
QY 841 AGCTTAAACCGCTGACTCACCGCGCGGGAAGTGAAGGAGGCTTAAGGAGGCGCTTC 900  
Db 1329128 AGCTTAAACCGCTGACTCACCGCGCGGGAAGTGAAGGAGGCTTAAGGAGGCGCTTC 1329069  
QY 901 GAAGCTCCTTCAATCAATGATTGAGAGCAAGTGTCCAGTTGAATGGGTTTCATGAAGCT 960  
Db 1329068 GAAGCTCCTTCAATCAATGATTGAGAGCAAGTGTCCAGTTGAATGGGTTTCATGAAGCT 1329009  
QY 961 ATATTAAACCATGTTAAGAACCAATCAATTTACTTAAGTACTTCCATAGTCAAGTGGT 1020  
Db 1329008 ATATTAAACCATGTTAAGAACCAATCAATTTACTTAAGTACTTCCATAGTCAAGTGGT 1328949  
QY 1021 GATCATGGAATCTTCAATTAAGTACTTGGGCGGCAAGTCTTTACTTCTCATCGG 1080  
Db 1328948 GATCATGGAATCTTCAATTAAGTACTTGGGCGGCAAGTCTTTACTTCTCATCGG 1328889  
QY 1081 ACCGCAGAACTACTGCTGATTAAACAAAGGAATTAAGCGGAGGACTCAATTCGGTTCT 1140  
Db 1328888 ACCGCAGAACTACTGCTGATTAAACAAAGGAATTAAGCGGAGGACTCAATTCGGTTCT 1328829  
QY 1141 TCTCGTGTGTTAATTTCTGAGCTCTTTTGTTCATCCCGCGCACCTTGGCGGTTGATCT 1200  
Db 1328828 TCTCGTGTGTTAATTTCTGAGCTCTTTTGTTCATCCCGCGCACCTTGGCGGTTGATCT 1328769  
QY 1201 TTTGTCCAAATCGCGCGCATGCTGCTGATATATGCGCTGGGCTGGCATCGCTTACCT 1260  
Db 1328768 TTTGTCCAAATCGCGCGCATGCTGCTGATATATGCGCTGGGCTGGCATCGCTTACCT 1328709  
QY 1261 GTTATGTTGCGGTCAATCGCAGCGAAGACGCCATGACAAACAAAGGTGAAGCGGCCACA 1320  
Db 1328708 GTTATGTTGCGGTCAATCGCAGCGAAGACGCCATGACAAACAAAGGTGAAGCGGCCACA 1328649  
QY 1321 GATCATTAAGAAACAGAACCAACCGTGGCCGATGACAGCTTTTGGCGGTTTCGCGCT 1380  
Db 1328648 GATCATTAAGAAACAGAACCAACCGTGGCCGATGACAGCTTTTGGCGGTTTCGCGCT 1328589  
QY 1381 GGCCTCATGACCGCAACCGGCTGGGCTGGAGGTGAGCGTGCATTAAGCAGCGGTTTG 1440  
Db 1328588 GGCCTCATGACCGCAACCGGCTGGGCTGGAGGTGAGCGTGCATTAAGCAGCGGTTTG 1328529  
QY 1441 GGTAAAGCCCATGTTGATGCGCAATCTGCTGACTGCTGTTGAACCCGATCGGTATTTGGA 1500  
Db 1328528 GGTAAAGCCCATGTTGATGCGCAATCTGCTGACTGCTGTTGAACCCGATCGGTATTTGGA 1328469  
QY 1501 CGGTTTGTGTTTATCGCGCGGCTCGCGCGCAATACGCGGACACCGGCGGTTT 1560

1328468	QCGCTTGTGTATCTCGCGCGCGTGGCGCGCAATACGGCGACACCGACGTCGATTTT	1328405
1561	QCGCGCTGGCGGTTTCGCGGCAAGCGCTGATCTGGTTCCCGCTGGTGGGTTTCGGCGACG	1520
1328408	CGCGCTCGGCGGTTTCGCGGCAAGCGCTGATCTGGTTCCCGCTGGTGGGTTTCGGCGACG	1328349
1621	AGCATTTGTCAGCGCGCTGTCAGCGCCCAAGGTTGGCGCTCGATCAAGCTGTCGTGGC	1680
1328348	AGCATTTGTCAGCGCGCTGTCAGCGCCCAAGGTTGGCGCTCGATCAAGCTGTCGTGGC	1328289
1681	AGTTCTGATCAACGCATTGGCCATCAAACTGATGTTGATGGTTAGTTCGCGGGTTTT	1740
1328288	AGTTGTCATGACCGCATTTGSCCATCAAACTGATGTTGATGGTTAGTTCGCGGGTTTT	1328229
1741	GGAAATCGGTGGCTTCGCGCMAATGTTGATCCGGCTGTCGGGAATCTCATCGATCGC	1800
1328228	GGAAATCGGTGGCTTCGCGCMAATGTTGATCCGGCTGTCGGGAATCTCATCGATCGC	1328169
1801	CTTCAACTCGSGCGTCAGAAAACTCAAGTTGTTGAGTGAATCAAGGCTGTGTCCAGCTG	1860
1328168	CTTCAACTCGSGCGTCAGAAAACTCAAGTTGTTGAGTGAATCAAGGCTGTGTCCAGCTG	1328109
1961	CTCACTCGAGAGACCAATCAATGCACTGTCACGGTATCGCGCGTACTCTCCCTTG	1920
1328108	CTCACTCGAGAGACCAATCAATGCACTGTCACGGTATCGCGCGTACTCTCCCTTG	1328049
1921	CTCGCGACACCCCATGCAAGGGCATCTGCGCAAGTGAATGCGCGCTTCTGGGCGAT	1980
1328048	CTCGCGACACCCCATGCAAGGGCATCTGCGCAAGTGAATGCGCGCTTCTGGGCGAT	1327989
1981	GTCACTGAGCTTGGCGACCAATCAATATTTGTTACGTTCAACATGCCCTCAGACAGGCA	2040
1327988	GTCACTGAGCTTGGCGACCAATCAATATTTGTTACGTTCAACATGCCCTCAGACAGGCA	1327929
2041	CTTACCTCTGGCTGGCGGGGAACCTCTCGAAATTCATCGAGATATTTGTCGTGACGAG	2100
1327928	CTTACCTCTGGCTGGCGGGGAACCTCTCGAAATTCATCGAGATATTTGTCGTGACGAG	1327869
2101	GCCCTCGCGCAAGTGGTGAGAAAGCAATGACCCAGACCATTTGGCAGCTGACTGCAG	2160
1327868	GCCCTCGCGCAAGTGGTGAGAAAGCAATGACCCAGACCATTTGGCAGCTGACTGCAG	1327809
2161	CAAGTTCTCAACCGTCATCGCCGGTTCTCTACCAAGATTAAATGATGAATGACTGTCG	2220
1327808	CAAGTTCTCAACCGTCATCGCCGGTTCTCTACCAAGATTAAATGATGAATGACTGTCG	1327749
2221	CTGATGAATCAGAAAGCGGCGAGCCCTCTCTCGCGCTGTAATCTCAGCGCCTTCGTGTGAG	2280
1327748	CTGATGAATCAGAAAGCGGCGAGCCCTCTCTCGCGCTGTAATCTCAGCGCCTTCGTGTGAG	1327689
2281	CTCTCGACCGTAGGAAGAAATACCCAAGTAAGAGGCTTTCCAGAGCAACAAATGTCACG	2340
1327688	CTCTCGACCGTAGGAAGAAATACCCAAGTAAGAGGCTTTCCAGAGCAACAAATGTCACG	1327628
2341	CAATCGCTACATGGTTCTTCCAAAGAGTATCT	2374
1327628	CAATCGCTACATGGTTCTTCCAAAGAGTATCT	1327595

RES-09-746-660A-47/C  
; Sequence 47, Application US/09746660A  
; Publication No. US20030049804A1  
; GENERAL INFORMATION:  
; APPLICANT: Pompeius, Markus  
; APPLICANT: Kroeger, Burkhard  
; APPLICANT: Schroder, Hartwig  
; APPLICANT: Zeldger, Onkar  
; APPLICANT: Habernauer, Gregor  
; APPLICANT: Kim, Jun-Mon  
; APPLICANT: Lee, Heung-Schick  
; APPLICANT: Hwang, Byung-Joon

Query Match	41.8%;	Score 993;	DB 10;	Length 993;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Gaps 0;
Matches 993;	Conservative	0;	Indels	0;
62	CTGAACCTTTTCAGAAAGTAACATAAGCGCCGCAATCCCTCGATTGCTGTGCATCAACGACGCGC	121		
993	CTGAACCTTTTCAGAAAGTAACATAAGCGCCGCAATCCCTCGATTGCTGTGCATCAACGACGCGC	934		
122	TCGTGTAGTCTAGCTAGAGATCTAGATTCCAGGGCGCCATCGTTGCCAATAATCATCGGTGTG	181		
913	TCGTGTAGTCTAGCTAGAGATCTAGATTCCAGGGCGCCATCGTTGCCAATAATCATCGGTGTG	874		
192	TCAAATGGGTATCATCGAGAGAGCATCACTTCTCTCTGTTTTCAGCATCGGAGCAGTGTGG	241		
873	TCAAATGGGTATCATCGAGAGAGCATCACTTCTCTCTGTTTTCAGCATCGGAGCAGTGTGG	814		
242	GTTTTCGGGAAGAAGTCCCCCAACCAAGGCTCTGGCGGAATTCGCTCACCAAAACCTTCGCGC	301		
813	GTTTTCGGGAAGAAGTCCCCCAACCAAGGCTCTGGCGGAATTCGCTCACCAAAACCTTCGCGC	754		
302	GACGGGACAAATGGATAGCGGCTGCGGCCACAGACCAATCGACGCGCCGCTCCAGAGTCA	361		
753	GACGGGACAAATGGATAGCGGCTGCGGCCACAGACCAATCGACGCGCCGCTCCAGAGTCA	694		
362	CGGTCCTTGAAGCACATCTTTTGGGACCGAAGCGCTAAGACGGGCGATCGAGGCCATCTAGT	421		
693	CGGTCCTTGAAGCACATCTTTTGGGACCGAAGCGCTAAGACGGGCGATCGAGGCCATCTAGT	634		
422	TTCCCATCAACCATGTAGGATCCCGCAGATGAGGGGGTTCGAAATGSCCAAGTGTGCGCGATG	481		
633	TTCCCATCAACCATGTAGGATCCCGCAGATGAGGGGGTTCGAAATGSCCAAGTGTGCGCGATG	574		
482	GTTCCAAGTTCTACTACTTCACTTCACTCCCGCCCA CGGGATTAGCTTCA CGGGTTACCGTTCCT	541		
573	GTTCCAAGTTCTACTACTTCACTTCACTCCCGCCCA CGGGATTAGCTTCA CGGGTTACCGTTCCT	514		
542	AAAAACATCTCCACGCGCAGCAAGATATGTGTGGGCTTCACTTCTTCCAGCGCAGCGGTG	601		
513	AAAAACATCTCCACGCGCAGCAAGATATGTGTGGGCTTCACTTCTTCCAGCGCAGCGGTG	454		
602	ACGCTTCTCCACCCCAAGAGACTACCTTCGTTGAAACGCGGAGGAAACCATGTGTGGATAGC	661		

810	AGATTCCAGGGGCATCGTTGCCAATACATCGGTGTGTCAATGGGTATCTCATCGAGGAG	751
205	GATCACTTCTCTCTGTTTTAGCATGGGAGCAGCTTGGGTTTCGGGAGAAGTCCCCACC	264
750	GATCACTTCTCTCTGTTTTAGCATGGGAGCAGCTTGGGTTTCGGGAGAAGTCCCCACC	691
265	AAGGCTTCGGGGAATTGCTCTCAACAACTTCCTCGCGACGGGCANTGATGATCGCGCT	324
690	AAGGCTTCGGGGAATTGCTCTCAACAACTTCCTCGCGACGGGCANTGATGATCGCGCT	631
325	GCGCCCCACAGGACCATGAGCGGCCCGCTCAGGTCACGGTCTTGAAGACACATCTTTGGG	384
630	GCGCCCCACAGGACCATGAGCGGCCCGCTCAGGTCACGGTCTTGAAGACACATCTTTGGG	571
385	ACCGAAGCGTAAAGACGGGCATCGCGACCCCAATCTAGTTTCCCATCAACCATGTAGGCATC	444
570	ACCGAAGCGTAAAGACGGGCATCGCGACCCCAATCTAGTTTCCCATCAACCATGTAGGCATC	511
445	CCGCAATGAGGGGGTTGCCAATGGCCAAAGTGGCGCATGGTTCCAGTTCTACTACTTCA	504
510	CCGCAATGAGGGGGTTGCCAATGGCCAAAGTGGCGCATGGTTCCAGTTCTACTACTTCA	451
505	TCCCGCGACGGGATTAGTCTCACGGGTTACGGTCTCTAAACCATCTCCACGGCGCAGCAA	564
450	TCCCGCGACGGGATTAGTCTCACGGGTTACGGTCTCTAAACCATCTCCACGGCGCAGCAA	391
565	GGATTAATGTGTGGGCTTCATCTTCCAAAGCGCAGCGTGAGCGTTGCTCCACCCCAAGAACG	624
390	GGATTAATGTGTGGGCTTCATCTTCCAAAGCGCAGCGTGAGCGTTGCTCCACCCCAAGAACG	331
625	TACCTCGTTGAACACGGGAGGAAACCATGTGGATAGCGAATCTGCGTTGATGGCGATGCT	684
330	TACCTCGTTGAACACGGGAGGAAACCATGTGGATAGCGAATCTGCGTTGATGGCGATGCT	271
685	TACGGGATTTACGAAAGCGGTCAGATAGTTTGGCTTTAGTTTCTGCTTGCAGCAACAC	744
270	TACGGGATTTACGAAAGCGGTCAGATAGTTTGGCTTTAGTTTCTGCTTGCAGCAACAC	211
745	CATTTTCCGCGCTGCTTGCAACAGGACTTACCCGCTTCGGTTGCTTTGGCCGGTTGGGT	804
210	CATTTTCCGCGCTGCTTGCAACAGGACTTACCCGCTTCGGTTGCTTTGGCCGGTTGGGT	151
805	GCGCATATCAACACTCGACCCACGTGATGCTCGAGAGCTTTAAACGCGCTGACTCACCGC	864
150	GCGCATATCAACACTCGACCCACGTGATGCTCGAGAGCTTTAAACGCGCTGACTCACCGC	91
865	CGAGGGGGAAATGGAAGGGGCTTAAGGAGGGCGCTTTCGAGAGCTGCTTCATCAATGATTGA	924
90	CGAGGGGGAAATGGAAGGGGCTTAAGGAGGGCGCTTTCGAGAGCTGCTTCATCAATGATTGA	31
925	GAGCAAGTGTCCAGTTTGAATGGGGTTTCAT	954
30	GAGCAAGTGTCCAGTTTGAATGGGGTTTCAT	1

 $Q_Y$ 

330	TACCTCGTTGAA	CACACGGAGGAAACCATGTGGATAGCGAATCTCGCTTGATGCGCATGGT	271
695	TACGGGATTT	CAGCAAGCGCTCCAGATAGTTGGCTTTAGTTTCTGCTTCGACGACAC	744
270	TAAACGGGATTT	CAGCAAGCGCTCCAGATAGTTGGCGCTTTAGTTTCTGCTTCGACGACAC	211
745	CATTTTCCGCGCT	GTTCGCACAAAGCACTTACCCGCTTCGGTTGCTTTGGCCGGTTGGGT	804
210	CATTTTCCGCGCT	GTTCGCACAAAGCACTTACCCGCTTCGGTTGCTTTGGCCGGTTGGGT	151
805	GCGCATACCA	CAACTCGACCCACGTGATGCTCGAGAGCTTTTAACGGCTGACTCACCGC	864
150	GCGCATACCA	CAACTCGACCCACGTGATGCTCGAGAGCTTTTAACGGCTGACTCACCGC	91
865	CGAGGGGGGAA	TGAAAGGGCTTAAGAGGGCGCTTCGAAGCTGCCTTCATCAATGATTGA	924
90	CGAGGGGGGAA	TGAAAGGGCTTAAGAGGGCGCTTCGAAGCTGCCTTCATCAATGATTGA	31
925	GAGCAAGTGTC	CAGTTGTAATGGGTTTCAT	954
30	GAGCAAGTGTC	CAGTTGTAATGGGTTTCAT	1

RESULT 5

US-09-746-660A-51

; Sequence 51, Application US/09746660A

; Publication No. US20030049804A1

; GENERAL INFORMATION:

; APPLICANT: Pompejus, Markus

; APPLICANT: Kroger, Burkhard

; APPLICANT: Schroder, Hartwig

; APPLICANT: Zelder, Oskar

; APPLICANT: Haberhauer, Gregor

; APPLICANT: Kim, Jun-Won

; APPLICANT: Lee, Heung-Schick

; APPLICANT: Hwang, Byung-Joon

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING

; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS

; FILE REFERENCE: BGI-121Cf2

; CURRENT APPLICATION NUMBER: US/09/746,660A

Thu Mar 18 12:30:59 2004

```
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patent In Vers. 2.0
; SEQ ID NO 51
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(799)
; OTHER INFORMATION: RXA01394
US-09-746-660A-51

Query Match      34.6%; Score 822; DB 10; Length 822;
Best Local Similarity 100.0%; Pred. No. 7.1e-264;
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 925 GAGCAAGTGCAGTTGAATCGGGTTCATGAAGCTATATTAACCATGTAAAGCAAA 984
DB 1 GAGCAAGTGCAGTTGAATCGGGTTCATGAAGCTATATTAACCATGTAAAGCAAA 60

QY 985 TCATTTTACTTAAGTACTTCCATAGTTCAGATGTCATGGAATCTTCAATACAGG 1044
DB 61 TCATTTTACTTAAGTACTTCCATAGTTCAGATGTCATGGAATCTTCAATACAGG 120

QY 1045 TCTGCTTTTGGGGCCAGTCTTTTACTGTCATCGACCGCGAATGACTGCTATTA 1104
DB 121 TCTGCTTTTGGGGCCAGTCTTTTACTGTCATCGACCGCGAATGACTGCTATTA 180

QY 1105 ACAAGAAATTAAGCGGAGGACTCATTTGGGTTCTTCTCGTGTGTTAATTTCTGAGT 1164
DB 181 ACAAGAAATTAAGCGGAGGACTCATTTGGGTTCTTCTCGTGTGTTAATTTCTGAGT 240

QY 1165 CTTTTGTTCATCGCGGACCTTTGGCGGTTTGATCTTTTGTCCAAATGCGCGCGATCGT 1224
DB 241 CTTTTGTTCATCGCGGACCTTTGGCGGTTTGATCTTTTGTCCAAATGCGCGCGATCGT 300

QY 1225 GCTCGATATTAATGCTGGGGTGGCATCGCTTACCTGTTATGCTTGGCTGATGCGAGC 1284
DB 301 GCTCGATATTAATGCTGGGGTGGCATCGCTTACCTGTTATGCTTGGCTGATGCGAGC 360

QY 1285 GAAAGAGCGCATGACAAACAGTGGAGCGCCACAGATCATTTGAAGAAACAGAACCAAC 1344
DB 361 GAAAGAGCGCATGACAAACAGTGGAGCGCCACAGATCATTTGAAGAAACAGAACCAAC 420

QY 1345 CGTGCCCGATGACAGCCCTTTGGCGGTTTCGGCGGTTGGCCACTGACACGCGCAACCGGT 1404
DB 421 CGTGCCCGATGACAGCCCTTTGGCGGTTTCGGCGGTTGGCCACTGACACGCGCAACCGGT 480

QY 1405 GCGGGTGGAGGTGAGCGTCGATTAAGCAGCGGGTTGGGTAAAGCCCATGTTGATGCGCAT 1464
DB 481 GCGGGTGGAGGTGAGCGTCGATTAAGCAGCGGGTTGGGTAAAGCCCATGTTGATGCGCAT 540

QY 1465 GGTGCTGACTGTTGAACCGAATGCTGTTTGAACCGGTTTGTGTTTATCGCGCGGT 1524
DB 541 GGTGCTGACTGTTGAACCGAATGCTGTTTGAACCGGTTTGTGTTTATCGCGCGGT 600

QY 1525 CGGCGCGCAATACGGCGACACCGGATGTTTTCGCGCTGGCGGTTCGCGGCAAG 1584
DB 601 CGGCGCGCAATACGGCGACACCGGATGTTTTCGCGCTGGCGGTTCGCGGCAAG 660
1585 CCGTATCTGTTCCCGCTGCTGGGTTTCGCGCGCAGCAGCATTTGTACGCGCGCTGTCCAG 1644
661 CCGTATCTGTTCCCGCTGCTGGGTTTCGCGCGCAGCAGCATTTGTACGCGCGCTGTCCAG 720
1645 CCCCCAAGGTGTGGCGCTGGATCAACAGTCGTCGTGGCAGTTGTGATACCGCATTTGGCCAT 1704
721 CCCCCAAGGTGTGGCGCTGGATCAACAGTCGTCGTGGCAGTTGTGATACCGCATTTGGCCAT 780
1705 CAAACTGATGTTGATGGGTTAGTTTTCGCGGGTTTTCGAATC 1746
781 CAAACTGATGTTGATGGGTTAGTTTTCGCGGGTTTTCGAATC 822

RESULT 6
US-10-166-142-7
; Sequence 7, Application US/10166142
; Publication No. US20030124687A1
; GENERAL INFORMATION:
; APPLICANT: YASUEDA, HISASHI
; APPLICANT: GUNJI, YOSHIYA
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL
; TITLE OF INVENTION: ASSIMILATING BACTERIUM
; FILE REFERENCE: 223789US
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: JP 2001-177075
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(711)
; OTHER INFORMATION:
US-10-166-142-7

Query Match      29.9%; Score 711; DB 14; Length 711;
Best Local Similarity 100.0%; Pred. No. 1e-226;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1016 ATGCTCATCATGAAATCTTCAATACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 1075
DB 1 ATGCTCATCATGAAATCTTCAATACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 60

QY 1076 ATCGGACCGCAGATGCTGCTGATTAACAAGAAATTAAGCGGAGGACTCATTTGCG 1135
DB 61 ATCGGACCGCAGATGCTGCTGATTAACAAGAAATTAAGCGGAGGACTCATTTGCG 120

QY 1136 GTTCTTCTGCTGTTTAAATTTCTGACGCTTTTGTTCATCGCGCGCACTTTGGCGGT 1195
DB 121 GTTCTTCTGCTGTTTAAATTTCTGACGCTTTTGTTCATCGCGCGCACTTTGGCGGT 180

QY 1196 GATCTTTTGTCCAAATGCGCGCGATCGTGTGCTGATATTAATGCTGGGTGCGATCGCT 1255
DB 181 GATCTTTTGTCCAAATGCGCGCGATCGTGTGCTGATATTAATGCTGGGTGCGATCGCT 240

QY 1256 TACCTGTATGTTTGGCGCTATGCGGAGGAGCGCATGACAAACAGGTGGAAGCG 1315
DB 241 TACCTGTATGTTTGGCGCTATGCGGAGGAGCGCATGACAAACAGGTGGAAGCG 300

QY 1316 CCACAGATCATTTGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTTGGCGGTTCG 1375
DB 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTTGGCGGTTCG 360

QY 1376 GCGGTGGCCACTGACACCGCAACCGGGTTCGCGGTGAGGTGAGGTGAGGTGAGGTGAGGT 1435
DB 361 GCGGTGGCCACTGACACCGCAACCGGGTTCGCGGTGAGGTGAGGTGAGGTGAGGTGAGGT 420
```

QY 1436 GTTGGGTAAGCCCATTTGATGGAATGCTGCTGATCTGTTGAACCGAATGCGTAT 1495  
 Db 421 GTTGGGTAAGCCCATTTGATGGAATGCTGCTGATCTGTTGAACCGAATGCGTAT 480  
 QY 1496 TTGGACGCTGTTGTTATCGCGCGCTCGCGCGCAATAGCGGACCGGCGGTGG 1555  
 Db 481 TTGGACGCTGTTGTTATCGCGCGCTCGCGCGCAATAGCGGACCGGCGGTGG 540  
 QY 1556 ATTTTCGCGCTGCGCGCTTCGCGCGCAAGCTGATCTGGTTCCCGCTGGTGGGTTTCGC 1615  
 Db 541 ATTTTCGCGCTGCGCGCTTCGCGCGCAAGCTGATCTGGTTCCCGCTGGTGGGTTTCGC 600  
 QY 1616 GCAGCAGCATTTGATGCGCGCTTCGCGCGCAAGCTGATCTGGTTCCCGCTGGTGGGTTTCGC 1675  
 Db 601 GCAGCAGCATTTGATGCGCGCTTCGCGCGCAAGCTGATCTGGTTCCCGCTGGTGGGTTTCGC 660  
 QY 1676 GTGCGAGTTGTCATGACCGCATTTGGCGCATCAAACTGATGTTGATGGGTTAG 1726  
 Db 661 GTGCGAGTTGTCATGACCGCATTTGGCGCATCAAACTGATGTTGATGGGTTAG 1711

RESULT 7  
 US-09-738-626-3455  
 ; Sequence 3455, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAOKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738, 626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR FILING DATE: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 3455  
 ; LENGTH: 708  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-3455

Query Match 29.8%; Score 708; DB 9; Length 708;  
 Best Local Similarity 100.0%; Pred. No. 1e-225;  
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1016 ATGGTGATCATGGAATCTTCAATACAGGTCGCTTTGGGGCCAGCTTTTACTGTCC 1075  
 Db 1 ATGGTGATCATGGAATCTTCAATACAGGTCGCTTTGGGGCCAGCTTTTACTGTCC 60  
 QY 1076 ATCGGACCGGAGATGCTGATTAACAGGAATTAAGCGGAGGACTCATTTGG 1135  
 Db 61 ATCGGACCGGAGATGCTGATTAACAGGAATTAAGCGGAGGACTCATTTGG 120  
 QY 1136 GTTCTCTCGTGTTTAAATTCGAGTCTTTTGTTCATCGCGGACCTTTGGCGTT 1195  
 Db 121 GTTCTCTCGTGTTTAAATTCGAGTCTTTTGTTCATCGCGGACCTTTGGCGTT 180  
 QY 1196 GATCTTTTGTCCAAATGCCGCGCATGCTGCTGATTAATGCGCTGGGCTGGCATCGCT 1255  
 Db 181 GATCTTTTGTCCAAATGCCGCGCATGCTGCTGATTAATGCGCTGGGCTGGCATCGCT 240

QY 1256 TACCTGTTATGTTTGGCTCATGCGGAGAAAGCGCCATGACAAACAGGTCGAGCG 1315  
 Db 241 TACCTGTTATGTTTGGCTCATGCGGAGAAAGCGCCATGACAAACAGGTCGAGCG 300  
 QY 1316 CCACAGATCATTTGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGCGGTTCC 1375  
 Db 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGCGGTTCC 360  
 QY 1376 CGCTGGCCATGACACCGGCAACCGGCTGGGCTGGGCTGAGGTCGATTAAGCAGCG 1435  
 Db 361 CGCTGGCCATGACACCGGCAACCGGCTGGGCTGGGCTGAGGTCGATTAAGCAGCG 420  
 QY 1436 GTTTCGGTAAAGCCCATGCTGATGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1495  
 Db 421 GTTTCGGTAAAGCCCATGCTGATGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 1496 TTGGACGCTGTTGTTTATCGCGCGCTCGCGCGCAATACCGCGACACCGGACGCTGG 1555  
 Db 481 TTGGACGCTGTTGTTTATCGCGCGCTCGCGCGCAATACCGCGACACCGGACGCTGG 540  
 QY 1556 ATTTTCGCGCTGGCGCGCTTCGCGCGCAAGCTGATCTGGTTCCCGCTGGTGGGTTTCGC 1615  
 Db 541 ATTTTCGCGCTGGCGCGCTTCGCGCGCAAGCTGATCTGGTTCCCGCTGGTGGGTTTCGC 600  
 QY 1616 GCAGCAGCATTTGTCAGCGCGCTGCTGCGCGCAAGGCTGGGCTGGATCAACGTCGTC 1675  
 Db 601 GCAGCAGCATTTGTCAGCGCGCTGCTGCGCGCAAGGCTGGGCTGGATCAACGTCGTC 660  
 QY 1676 GTGCGAGTTGTCATGACCGCATTTGGCGCATCAAACTGATGTTGATGGGTT 1723  
 Db 661 GTGCGAGTTGTCATGACCGCATTTGGCGCATCAAACTGATGTTGATGGGTT 708

RESULT 8  
 US-10-166-142-9  
 ; Sequence 9, Application US/10166142  
 ; Publication No. US20030124687A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GUNJI, YOSHIYA  
 ; APPLICANT: YASUEDA, HISASHI  
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL  
 ; TITLE OF INVENTION: ASSIMILATING BACTERIUM  
 ; FILE REFERENCE: 223789US  
 ; CURRENT APPLICATION NUMBER: US/10/166,142  
 ; CURRENT FILING DATE: 2002-06-11  
 ; PRIOR APPLICATION NUMBER: JP 2001-1777075  
 ; PRIOR FILING DATE: 2001-06-12  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 9  
 ; LENGTH: 712  
 ; TYPE: DNA  
 ; ORGANISM: Brevibacterium lactofermentum  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(375)  
 ; OTHER INFORMATION:  
 US-10-166-142-9

Query Match 29.4%; Score 696.8; DB 14; Length 712;  
 Best Local Similarity 99.6%; Pred. No. 5.7e-222;  
 Matches 709; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 QY 1016 ATGGTGATCATGGAATCTTCAATACAGGTCGCTTTGGGGCCAGCTTTTACTGTCC 1075  
 Db 1 ATGGTGATCATGGAATCTTCAATACAGGTCGCTTTGGGGCCAGCTTTTACTGTCC 60  
 QY 1076 ATCGGACCGGAGATGCTGATTAACAGGAATTAAGCGGAGGACTCATTTGG 1135  
 Db 61 ATCGGACCGGAGATGCTGATTAACAGGAATTAAGCGGAGGACTCATTTGG 120  
 QY 1136 GTTCTCTCGTGTTTAAATTCGAGTCTTTTGTTCATCGCGGACCTTTGGCGTT 1195

Db 121 GTTCTTCTCGTGTGTTAAATTTCTGACGTCCTTTTGTTCATCGCGGACCTTGGGGGT 180  
Qy 1196 GATCTTTTGTCCAAATGCGCGCGGATCGTGTCTCGATATTATGCGCTGGGGTGGCATCGT 1255  
Db 181 GATCTTTTGTCCAAATGCGCGCGGATCGTGTCTCGATATTATGCGCTGGGGTGGCATCGT 240  
Qy 1256 TACCTGTATGTTTGGCGTATGCGGAGGAAAGAGCCCATGACAAACAAGGTGGAGCG 1315  
Db 241 TACCTGTATGTTTGGCGTATGCGGAGGAAAGAGCCCATGACAAACAAGGTGGAGCG 300  
Qy 1316 CCACAGATCATTTGAAGAAACAGAACCAACCGTCCCGATGACACGCTTTTGGCG-GTTC 1374  
Db 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTCCCGATGACACGCTTTTGGCGGTTC 360  
Qy 1375 GCGGCTGGCCACTGACACGCGCAACCGGGTGGGGTGGAGGTGAGGTGATGAGAGCG 1434  
Db 361 GCGGCTGGCCACTGACACGCGCAACCGGGTGGGGTGGAGGTGAGGTGATGAGAGCG 420  
Qy 1435 GGTGTGGGTAAAGCCCATGTTGATGCGGCGGCGTCCCGCGCAATACGGCGCACCGGACGGTG 1494  
Db 421 GGTGTGGGTAAAGCCCATGTTGATGCGGCGGCGTCCCGCGCAATACGGCGCACCGGAGTGGTA 480  
Qy 1495 TTTTGACGCGTGTGTTTATCGCGGCGGCGTCCCGCGCAATACGGCGCACCGGACGGTG 1554  
Db 481 TTTTGACGCGTGTGTTTATCGCGGCGGCGTCCCGCGCAATACGGCGCACCGGACGGTG 540  
Qy 1555 GATTTTCCGCGTGGCGGCTTTCGCGCAAGCCCTGATCTGGTCCCGCTGGTGGTTCGG 1614  
Db 541 GATTTTCCGCGTGGCGGCTTTCGCGCAAGCCCTGATCTGGTCCCGCTGGTGGTTCGG 600  
Qy 1615 GCGAGCAGATTTGACGCGGCGTTCGAGCCCGAGGTGGCGTGGATCAACGTCGT 1674  
Db 601 GCGAGCAGATTTGACGCGGCGTTCGAGCCCGAGGTGGCGTGGATCAACGTCGT 660  
Qy 1675 COTGGCATGTGATGACCGCATTTGGCCATFCAAACTGATGTTGATGGGTAG 1726  
Db 661 COTGGCATGTGATGACCGCATTTGGCCATFCAAACTGATGTTGATGGGTAG 712

RESULT 9  
US-09-738-626-3454/c  
; Sequence 3454, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 3454  
; LENGTH: 627  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3454

Query Match 26.4%; Score 627; DB 9; Length 627;  
Best Local Similarity 100.0%; Pred. No. 1.3e-198;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1726 GTTTTCGCGGTTTTGGAAATCGGTGGCTTCGCCAAATGTTGATCGCGGCTGGTGGGA 1785  
Db 627 GTTTTCGCGGTTTTGGAAATCGGTGGCTTCGCCAAATGTTGATCGCGGCTGGTGGGA 568  
Qy 1786 AATCTCATGATCGCTCCAACTCGGCGTCAGAAAACCTCCAAAGTGTGTGAGTGAATCAAG 1845  
Db 567 AATCTCATGATCGCTCCAACTCGGCGTCAGAAAACCTCCAAAGTGTGTGAGTGAATCAAG 508  
Qy 1846 GCTGTGTTCAGTGTCTCAACTGACGAAGCACAATCAATGACTGCTCACGGTATCGCG 1905  
Db 507 GCTGTGTTCAGTGTCTCAACTGACGAAGCACAATCAATGACTGCTCACGGTATCGCG 448  
Qy 1906 GCGGTACTCTCTTGTCTGGCGCAGCACCCATCGCAAGCGCATCTGCGCAAGTGAATGCGCC 1965  
Db 447 GCGGTACTCTCTTGTCTGGCGCAGCACCCATCGCAAGCGCATCTGCGCAAGTGAATGCGCC 388  
Qy 1966 GGTTCCTCGGCGCATGTCAATTGAGCTTGGCGACCATATCAATATTGTTCACTTCAACAT 2025  
Db 387 GGTTCCTCGGCGCATGTCAATTGAGCTTGGCGACCATATCAATATTGTTCACTTCAACAT 328  
Qy 2026 GCGCTCAGACAGGACTTACCTCTGGCTGGCGCGGGAACCCCTCTGGAATTCATCGAGATA 2085  
Db 327 GCGCTCAGACAGGACTTACCTCTGGCTGGCGCGGGAACCCCTCTGGAATTCATCGAGATA 268  
Qy 2086 TTTGTCCGTGAGCAGCGCCCTGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCAATTGTT 2145  
Db 267 TTTGTCCGTGAGCAGCGCCCTGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCAATTGTT 208  
Qy 2146 GGCAGCTGACTGCAACAAAGTCTCACCGTTCATCGCCGCTTCCTCCACCAAGCAATTAAT 2205  
Db 207 GGCAGCTGACTGCAACAAAGTCTCACCGTTCATCGCCGCTTCCTCCACCAAGCAATTAAT 148  
Qy 2206 GATGAATAGCTTGGCTGATGAATCAGAGCGGGGAGCCCTCTCCGCCATGAATCTCAGC 2265  
Db 147 GATGAATAGCTTGGCTGATGAATCAGAGCGGGGAGCCCTCTCCGCCATGAATCTCAGC 88  
Qy 2266 GCGCTCCGCTGTGAGCTCTGGACCGTAGAGAAATACCCAGTAAAGAGCCCTTCCAGA 2325  
Db 87 GCGCTCCGCTGTGAGCTCTGGACCGTAGAGAAATACCCAGTAAAGAGCCCTTCCAGA 28  
Qy 2326 CGCAACAATGTACGCAATCGCTACAT 2352  
Db 27 CGCAACAATGTACGCAATCGCTACAT 1

RESULT 10  
US-10-627-476-669/c  
; Sequence 669, Application US/10627476  
; Publication No. US20040030116A1  
; GENERAL INFORMATION:  
; APPLICANT: Pompejue, Mark  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schoder, Hartwig  
; APPLICANT: Zeider, Oskar  
; APPLICANT: Haberhauser, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE  
; TITLE OF INVENTION: TRANSPORT  
; FILE REFERENCE: BGI-125CPCN  
; CURRENT APPLICATION NUMBER: US/10/627,476  
; CURRENT FILING DATE: 2003-07-25  
; PRIOR APPLICATION NUMBER: 09/602,787  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: USSN 60/141031  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: DE 19931454.3  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931478.0  
; PRIOR FILING DATE: 1999-07-08

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl

; AFFIDICANT: MALONE, CHERYL

675

•



QY 2150 GCTGACTGCAACAAAGTTCTCACCGTCATCGCCCGGTTCTCCACCCAAAGATTAAATGATG 2209  
Db 674 GCAGACTCTTAACAGGTTCTCGCCATCTTCACTGTTCTTTACCCAGCGGTTCAAGATA 615  
QY 2210 GAATAGCTTGGTGATCAATACAGAGCGGCGAGCCCTCTCCGCGCATGMACTCAGCCGCC 2269  
Db 614 GAATAGCTCGGTTGATGAATCAGAGCGGCGAGCCCTCTCCAGCTAGATTTCCGAGATC 555  
QY 2270 TCCGCTGTGAGCTCTTGACCGTAGGAGAAATACCCACGTAAGAGCCCTTCCAGACGCA 2329  
Db 554 TCTGCGGTTAGTTCCAGTCCATAGGAGAAATTCACATAGCAGGCTTGGCGGATGCC 495  
QY 2330 ACAAATGACCAAGTGGGTACATGTTCTTCCAAAGGAGTATCT 2374  
Db 494 ACGATGTCTCGAAAGCATATGCAAGTTCTTCCAAAGGCGGTATCT 450

RESULT 12  
US-10-282-122A-17523/c  
; Sequence 17523, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA 034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; PRIOR FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17523  
; LENGTH: 879  
; TYPE: DNA  
; ORGANISM: *Corynebacterium diptheriae*  
US-10-282-122A-17523

Query Match 12.8%; Score 302.8; DB 12; Length 879;  
Best local similarity 59.2%; Pred. No. 7e-90;  
Matches 517; Conservative 0; Mismatches 357; Indels 0; Gaps 0;  
QY 81 ACTAAGCGCGCAATCCCTCGATTGCTGCATCAACGAGCGGCTCTGTGAGTCTAGTAGAG 140

Db 874 ACAGTGCATGATAGTCCCTCGAGGCGTGCATCCACTACCGCTTGAAGTCAACCGCTTAGGG 815  
QY 141 ATCTAGATTCCAGGCGGCATCGTTGCCAATATACCGGTGTGTCAATGGGTATCTCATCGA 200  
Db 814 CTGAGATTTCACACCGCCATCGCTGCCAATATACAGCGAACTTCGTGCACCTCGCTATCTA 755  
QY 201 GGAGGATCACTTCTCTCTGCTTTTATGATGAGGAGCAGCTTGGGTTTCGGGAGAGAGTCCCC 260  
Db 754 GCATCACTACCTCACCAAGATCCAAAGCGGTAAAGCTTGAATATCGGGGAGCAATGCCCC 695  
QY 261 AACAGGCTCTCGGCAATTCCTACCAAAACCTTCGCGCCGAGCGGCAATGATACAGC 320  
Db 694 AGCCCAATCCACACGCGCGCTTCCATGATGCTCCGATGAAGGATCTCATTAACAC 635  
QY 321 GCCTGGCGCCCAACAGGACCATCGACGCGCCCTCCAGGTCAAGCTTTGAAGGACATCTT 380  
Db 634 GTCGCCCTGCGGAACTTCCACCAAGAGAGCGTTAAATATCAAGCTTCTGAAGCCCATCCC 575  
QY 381 TGGGACCGAAGAGGTAAAGACGGGCAATCGAGCCCAATCTAGTTTCCCATCAACCATGTAGG 440  
Db 574 GAGGGCCCAATCGCAATATCTGCGATCCGCTCCCAATCAACAGACCGCTTATGGTGTAGC 515  
QY 441 CATCCGCAATGAGGGGCTTGCATATGCGCATGCTTCCAAAGTTCCTACTACTT 500  
Db 514 GATCCAAACAGCTTTCGCACTTGCCTGCAAAATATCGAATGCACTTAAGCTATCGAGT 455  
QY 501 CACATCCCGCCACGCGGATTAGCTTTCAGGGGTTACCGCTTCTTAAAAATCTCCAGCGCGCA 560  
Db 454 CGCAACAGATACAGGTCTCTTTTCCCGGTTACAGCGCCAAAGACATCTCCGCGAGCA 395  
QY 561 GCAAGGATAATGTGCGCTTCACTTCCAGCGGAGGTGAGCGGTGCTCCACCCCAAG 620  
Db 394 ACACGCAACAGAAATGCGACTCTCTCGATACGATATGCAATGTGGCGGTATCACTCT 335  
QY 621 AAGCTACTCTGTTGACAGCGGAGGAAACCATGTGATAGCGAATCTGCGTTGATGCGCA 680  
Db 334 CGCGCCACACTTGCCTATCACCGCGGAAACCATGTGCGCAAGATCTGCGTTGACAGTA 275  
QY 681 TGGTTAAACGCGGATTTTCAGCAAGCGCTCCAGATAGTTTGGCGTTTATGTTTCTGCTTCAGCA 740  
Db 274 CAGACATGGAATCCGCTCAATTTCTCCCGCAACTGAGCGTTTGTTCGGCTCGAGAA 215  
QY 741 ACACATTTTCGCGCTGCTTGCACAGGACTTCAACCGCTTCCGTTTGGGCGGCTT 800  
Db 214 GTGCCATACGTCTGCGGATTCACAGTACCTTCGCGCGGATCAGTTGACAGTGA 155  
QY 801 GGGTGGCGATACCAACACTCGACCCAGTGTCTCGAGAGCTTTAAACGCGCTGACTCA 860  
Db 154 AGTACGCGGTACCACTACTCGACGCTTCCCGCTCCAAAGCTTAATCTCTGACTTA 95  
QY 861 CCGCGGAGGCGGAAATGGAAGGGCTTAAGAGGCGCTTCCAGGCGCTTCCAGTCTCATCAATGA 920  
Db 94 CTGCTGAAGAGGAAATTCCTAATGCAAAAGCGCATCTTCGAAGCTGCTTCTGCGATTA 35  
QY 921 TTGAGAGCAAGTGTCCAGTGTGAATGGGTTCAAT 954  
Db 34 TTGCGAGCAAGGTTTCTTAATGAAGCGGATTCAT 1

RESULT 13  
US-10-369-493-35374/c  
; Sequence 35374, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 35374  
 ; LENGTH: 1026  
 ; TYPE: DNA  
 ; ORGANISM: Agrobacterium tumefaciens  
 US-10-369-493-35374

Query Match 6.9%; Score 164.4; DB 15; Length 1026;  
 Best Local Similarity 57.2%; Pred. No. 1.8e-43;  
 Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;

QY	1745	TCGGTGGCTTCGCGCAAAATGTTGATGCGCGGTGCGGAATCTCATCGATCGCTCC	1804
DB	1022	TCGGCAGATTTTGGCCAGAGTTGATATCCGATCTTGGCGTAACGGTTCGATTTGGGCC	963
QY	1805	AACCTCGCGCTCAGAAAATCTCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGCTGTCTCA	1864
DB	962	AGCTCTCTCGGTAGAGAACTCGGCAATTCGAGTGTCTTACGAGTCTCTCGACCTGTTCG	903
QY	1865	ACTGACGAAGCAACCAATCAATGCACTGTGTCAGGTATCCGCGCGGTACTCTCTTGTCTCG	1924
DB	902	ACACGGCTTGGCCCAATCAATGCTGAGGTAATGCGGCCG-----CG	861
QY	1925	CGCAGACCCATGCGAGCGCATCTGCGAAGTGAATGCGCGGTTCCTGGCGCATGTCA	1984
DB	860	CGCAGAACCCGCGCAATTTGCCATCTGCGCAGCGTCTGCGCAAGCGCTCGGCAATGCTG	801
QY	1985	TTGAGCTTGGCGACCAATCAATATTTTTCAGTTTCAATGTTTCAATGTTTCAATGTTTCA	2044
DB	800	TTGAGCGCGCGATGTTTTCGACATTTGCGCTCTGTTGAGGAAGCGCGGTGAGTGACTTG	741
QY	2045	CCCTGGTGGCGGGAACCTCTGGAATTCATCGAGATTTTGTTCGTCGAGCAGCGCC	2104
DB	740	CTCTGTGAGCAGCGCTGCCATCGGCACACCGCCCGAGATTTTGTCTGTCAGCATGCC	681
QY	2105	TGCGCAAGTGTGAGAAAGCAATGACCGCAAGCAATTTGTTGGCAGCTGACTGCAACAG	2164
DB	680	TGCGCCAGCGCGGAAGAGCATGAGAGCGCATACCGAGTTCTTCCAGCGTAT-----	629
QY	2165	TTCTCACCGTCTATCGCGGTCTCTCCACCAAGATTAATGATGATGATGATGATGATGATG	2224
DB	628	-----CGACAGACCGTCTCTCGATCGAGCGTTGATCATCGATAGCTCGGTG	576
QY	2225	TGAATCAGAGCGGCGAGCCCTCTCGGCATGAACTCAGCGCCCTTCGCTGTGAGCTCT	2284
DB	575	TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	516
QY	2285	GGACCGTAGGAGAAATACCCAGTAAGAGCTTTCCAGACGCAACATGTCACGCAAT	2344
DB	515	GAGTTGTAGAGAGATGCCAGATAGAGCGCTTCCGAGCGCGCTTCGCGGTGCGCTTC	456
QY	2345	GGTACATGTTTCTTCCAAAGGAGTAT	2372
DB	455	GGCGCGCAGGTTTCTTCAAGCGGTGTGT	428

RESULT 14  
 US-10-369-493-38826/c  
 ; Sequence 38826, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 38826  
 ; LENGTH: 1026  
 ; TYPE: DNA  
 ; ORGANISM: Agrobacterium tumefaciens  
 US-10-369-493-38826

Query Match 6.9%; Score 164.4; DB 15; Length 1026;  
 Best Local Similarity 57.2%; Pred. No. 1.8e-43;  
 Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;

QY	1745	TCGGTGGCTTCGCGCAAAATGTTGATGCGCGGTGCGGAATCTCATCGATCGCTCC	1804
DB	1022	TCGGCAGATTTTGGCCAGAGTTGATATCCGATCTTGGCGTAACGGTTCGATTTGGGCC	963
QY	1805	AACCTCGCGCTCAGAAAATCTCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGCTGTCTCA	1864
DB	962	AGCTCTCTCGGTAGAGAACTCGGCAATTCGAGTGTCTTACGAGTCTCTCGACCTGTTCG	903
QY	1865	ACTGACGAAGCAACCAATCAATGCACTGTGTCAGGTATCCGCGCGGTACTCTCTTGTCTCG	1924
DB	902	ACACGGCTTGGCCCAATCAATGCTGAGGTAATGCGGCCG-----CG	861
QY	1925	CGCAGACCCATGCGAGCGCATCTGCGAAGTGAATGCGCGGTTCCTGGCGCATGTCA	1984
DB	860	CGCAGAACCCGCGCAATTTGCCATCTGCGCAGCGTCTGCGCAAGCGCTCGGCAATGCTG	801
QY	1985	TTGAGCTTGGCGACCAATCAATATTTTTCAGTTTCAATGTTTCAATGTTTCAATGTTTCA	2044
DB	800	TTGAGCGCGCGATGTTTTCGACATTTGCGCTCTGTTGAGGAAGCGCGGTGAGTGACTTG	741
QY	2045	CCCTGGTGGCGGGAACCTCTGGAATTCATCGAGATTTTGTTCGTCGAGCAGCGCC	2104
DB	740	CTCTGTGAGCAGCGCTGCCATCGGCACACCGCCCGAGATTTTGTCTGTCAGCATGCC	681
QY	2105	TGCGCAAGTGTGAGAAAGCAATGACCGCAAGCAATTTGTTGGCAGCTGACTGCAACAG	2164
DB	680	TGCGCCAGCGCGGAAGAGCATGAGAGCGCATACCGAGTTCTTCCAGCGTAT-----	629
QY	2165	TTCTCACCGTCTATCGCGGTCTCTCCACCAAGATTAATGATGATGATGATGATGATGATG	2224
DB	628	-----CGACAGACCGTCTCTCGATCGAGCGTTGATCATCGATAGCTCGGTG	576
QY	2225	TGAATCAGAGCGGCGAGCCCTCTCGGCATGAACTCAGCGCCCTTCGCTGTGAGCTCT	2284
DB	575	TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	516
QY	2285	GGACCGTAGGAGAAATACCCAGTAAGAGCTTTCCAGACGCAACATGTCACGCAAT	2344
DB	515	GAGTTGTAGAGAGATGCCAGATAGAGCGCTTCCGAGCGCGCTTCGCGGTGCGCTTC	456
QY	2345	GGTACATGTTTCTTCCAAAGGAGTAT	2372
DB	455	GGCGCGCAGGTTTCTTCAAGCGGTGTGT	428

RESULT 15  
 US-10-369-493-38124/c  
 ; Sequence 38124, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B

Search completed: March 17, 2004, 09:46:30  
Job time : 870.071 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:43:00 ; Search time 6528.5 Seconds  
(without alignments)  
10858.975 Million cell updates/sec

Title: US-09-105-117K-1  
Perfect score: 2374  
Sequence: 1 ccattgtggaaggtttac.....ttttccaaaggatattc 2374

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthm:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pla:\*  
20: em\_gss\_vrc:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_png:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84.6	3.6	657	13 BQ155035	BQ155035 NF075E081
2	73.2	3.1	3237	28 BH770958	BH770958 LIMGTAG68
3	72	3.0	860	28 AF075981	AF075981 AF075981
4	60.4	2.5	830	29 CNS01MQH	AL151258 Anopheles

5	56.8	2.4	1034	28	B2554294
6	54.2	2.3	897	28	B2560550
7	51.4	2.2	1620	28	B2568946
8	48.8	2.1	1101	29	CNS017SY
9	48.6	2.0	833	28	B2556963
10	48.6	2.0	925	29	CNS0091P
11	47.6	2.0	852	13	BX333687
12	46	1.9	614	14	CA830247
13	46	1.9	720	14	CD881613
14	44.4	1.9	620	14	CA829022
15	44.4	1.9	644	13	BQ704105
16	44.4	1.9	664	29	CNS03K4J
17	44.4	1.9	1201	13	BX381961
18	43.6	1.8	721	28	B2562255
19	43.2	1.8	1395	28	B2574094
20	42.8	1.8	532	10	AW927407
21	42.8	1.8	538	14	CD861975
22	42.4	1.8	699	13	BX424825
23	42.2	1.8	885	13	BX425603
24	42	1.8	457	13	BY246241
25	41.8	1.8	1201	13	BX381961
26	41.8	1.8	947	29	CNS077MV
27	41.8	1.8	1083	29	CNS079WN
28	41.6	1.8	432	14	CB639421
29	41.6	1.8	558	12	BM712036
30	41.6	1.8	588	13	BU220119
31	41.6	1.8	772	14	CB668778
32	41.6	1.8	803	14	CB669972
33	41.6	1.8	813	14	CB684812
34	41.6	1.8	1201	13	BX356664
35	41.6	1.8	1201	13	BX361080
36	41.2	1.7	910	29	CNS0060N
37	41	1.7	617	12	BM729055
38	41	1.7	770	10	BE989495
39	41	1.7	861	10	BE740880
40	40.8	1.7	605	13	CA127255
41	40.8	1.7	614	13	CA141725
42	40.8	1.7	619	14	CA192339
43	40.8	1.7	619	14	CA238619
44	40.8	1.7	950	11	AY104298
45	40.6	1.7	436	28	AQ405681

## ALIGNMENTS

RESULT 1  
BQ155035  
LOCUS NF075E081F1067 Irradiated Medicago truncatula cDNA clone  
DEFINITION NF075E081R 5', mRNA sequence.  
ACCESSION BQ155035  
VERSION BQ155035.1 GI:20292094  
KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyle;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.  
REFERENCE 1 (bases 1 to 657)  
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,  
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula irradiated library  
Unpublished (2001)  
CONTACT: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org

Insert Length: 657 Std Error: 0.00  
 Plate: 075 row: E column: 08  
 Seq primer: TCACAGGAAACAGCTATGAC.  
 Location/Qualifiers  
 1. .657  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3880"  
 /clone="NF075E08IR"  
 /tissue\_type="seedlings"  
 /dev\_stage="seedling"  
 /clone\_lib="Irradiated"  
 /note="Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

ORIGIN

Query Match 3.6%; Score 84.6; DB 13; Length 657;  
 Best Local Similarity 48.2%; Pred. No. 3.4e-11;  
 Matches 237; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

464 ATGCCGAGTGGCGATGTTCCAAAGTTCTACTTCCATCCCGCCACGCGATTAGCT 523  
 Db  
 114 AGAACAGATAGTCAGCGCCACCAAGTTATCGACAGCAACTCCGCGAGCGCTGATG 173  
 524 TCACGGGTTACCGCTCTAAACATCTCCACGCGCAGCAGGATTAATGTGCGCTCA 583  
 Db  
 174 TGNATATCACCGCGCGACCACTTCGCGCGCGCAGACGTTCTCTGAGTGGCGTTCA 233  
 584 TCTTCAGCGCAGCGTGGCGTTGTCCTCACCCCAAGAGCTACCTCGTTGAACACGGA 643  
 Db  
 234 TCTTCTACCTGCAAGTTGAGCGGATAGCGGAATCAGCCACACAGGACCGAGTGCAG 293  
 644 GGAACCATCTGTATGATGATCTGCGTGTATGGGATGTTAAGCGGATTTACGACAG 703  
 Db  
 294 AGCAACACGTCGCGCAGCTGTCGCGGTGACCGCGAGTGAAGCAGCAGCGAGTGA 353  
 704 CGTCAGATAGTTGCGCTTTAGTTTCTGTTGAGCAACACCAATTTTCGCGCTGTGC 763  
 Db  
 354 CCGGTTTGTTCATCGCCCGACGCACTCTTCTTCAGCAACTCCACCTGGCGCAGTGC 413  
 764 ACAAGACTTCACCGCTTCGTTGCTTTGCGCGGTGGTGGCGATACCAACTCGA 823  
 Db  
 414 AGCAGTTTGTTCGCTTTGTTTCGTCGCGCGCGGTAGCGTACGACCAACAGCGCTGC 473  
 824 CCACGTCATCTCGAGAGCTTTAAAGCGCTGACTCACCGCGAGGGGGAATGGAAGG 883  
 Db  
 474 CGAACATATTTTCCAGTTGCTTAAATGCGCTGTGAGACGCGTGTATGATGACAGC 533  
 884 GCTAAGAGGCGCGCTTCGAAGCTGCTTCAATGATGATGAGACAAAGTTCAGTTGA 943  
 Db  
 534 TTTTGTGCGCGCGCTCAATCTCTGTTACGATATCACCGATCCAGTGCCTGTAATGTT 593  
 944 ATGGGGTTCATG 955  
 594 CTGTAGTCCNGG 605

RESULT 2

BH770958 3237 bp DNA linear GSS 01-MAY-2002  
 LOCUS  
 DEFINITION LLMGtag686 MG1363 Random Sequence Tag Library Lactococcus lactis

FEATURES

subsp. cremoris genomic, genomic survey sequence.  
 BH770958  
 VERSION BH770958.1 GI:20373915  
 KEYWORDS GSS.  
 SOURCE Lactococcus lactis subsp. cremoris  
 ORGANISM Lactococcus lactis subsp. cremoris  
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
 REFERENCE 1. (bases 1 to 3237)  
 Bolotin, A., Ehrlich, S.D. and Sorokin, A.  
 Studies of genomes of dairy bacteria Lactococcus lactis  
 TITLE Sci. Aliments (2002) In press  
 JOURNAL Contact: Sorokin A  
 COMMENT Genetique Microbienn  
 INRA  
 CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France  
 Tel: 33 1 34 65 25 16  
 Fax: 33 1 34 65 25 21  
 Email: sorokine@jouy.inra.fr  
 best homologue in strain IL1403 is nadR (98%)  
 Class: shotgun  
 High quality sequence start: 30  
 High quality sequence stop: 3207.  
 FEATURES  
 source  
 1. .3237  
 /organism="Lactococcus lactis subsp. cremoris"  
 /mol\_type="genomic DNA"  
 /strain="MG1363"  
 /sub\_species="cremoris"  
 /db\_xref="taxon:1359"  
 /clone\_lib="MG1363 Random Sequence Tag Library"  
 /notes="Vector: pSGM02; Site 1: SmaI; Library of chromosomal fragments of L.lactis strain MG1363 was prepared by partial AluI digestion or by sonication."

ORIGIN

Query Match 3.1%; Score 73.2; DB 28; Length 3237;  
 Best Local Similarity 49.8%; Pred. No. 9.5e-08;  
 Matches 287; Conservative 0; Mismatches 263; Indels 26; Gaps 3;

1805 AACTCGGCGTCAGAAATCCAGTTCTTCAGTGAATCAAGCTGTGTCCAGCTGCTCA 1864  
 QY  
 1956 AATTCTTCGTGAGTAATTCAGAGTTCTTAAGCTGCAATTTTCATGATGTTGTGC 2015  
 Db  
 1865 ACTGACGAAGCACAATCAATGACATCGTTCACGTTATCGGCGCTACTCTCTGCTGC 1924  
 QY  
 2016 GGACGACTTGGCCCAATCAAGCTGATGAACAGCTTGAACCTTTGCCAGCTCTT---TCC 2072  
 Db  
 1925 CGCAGACCCATGACAGCCCATCTGCGGAGTACTGCGCGCTTCTGGGCGATGTCA 1984  
 QY  
 2073 CGTAAACCCCAAGCTTATGCCATTGAGCTAGAGATTGTCCACGACTTTGAGCTAAATCA 2132  
 Db  
 1985 TTGAGCTTGGGACATATCAATATTTTCAGTTTCAATGCTTCAACATGCCCTCAGACAGGACTTA 2044  
 QY  
 2133 TTAAGGCTTGAATGTTGTTCCATCTTTCTGTTGTAAGACTGTCAATCATGCAAGTCA 2192  
 Db  
 2045 CCCTGCTGGCGCGG-----GAACCTCTGGAATTCATCGAGATATTGTCCGTGAGC 2098  
 QY  
 2193 TAATGAGGATCAGCAATTCGCAATTTTCAGGAATTTCCATGCAATATATTTCTGTTAAG 2252  
 Db  
 2099 AGGCCCTGCGCAAGTGTGTGAGAACATGACCGAGACCATTTGTTGGCAGCTGACTGC 2158  
 QY  
 2253 AGTCTCTGATACAAAGGCTTTAAAGGCAATTTGTTCTTATCCC-----CC 2295  
 Db  
 2159 AACAGTTCTTCAACGCTTCATCGCCGCTTCTCCACCAACGATTAATGATGAATAGCTT 2218  
 QY  
 2296 CCCTCTGTCAAAGTTCTTGTAAATCATCTTCAATCCATCGATCTAACATTGAATACGA 2355  
 Db  
 2219 GGCTGATGAATCAGAGCGGCGAGCCCTCTCTCGCGCATGAACTCAGCGGCTCCGCTGTG 2278  
 QY  
 2356 GGCTGATGAATCAAGAGTTTAAAGGCTTAACTTTTCAGCGCGCTTAAGACCGGAGCTTGGTT 2415  
 Db  
 2279 AGCTCTGACCGTAGGAGGAATATCCACGTTAAGAGGCTTTCCAGAGCGCAACAATGTCA 2338  
 QY

Db 2416 TCTTGGGAGAAATATTAGACAAACCCATAAAGTGCTTTTCCATATCGAGCAGTT 2475  
 QY 2339 CGCAATGCGTACATGGTTTCTTCCAAAGGAGTATCT 2374  
 Db 2476 TCCAAAGCGCCCATGGTTTCTTCAATAGGGGTATT 2511

RESULT 3  
 AF075981/c  
 LOCUS  
 DEFINITION AF075981 Salmonella typhimurium LT2, Lambda DASH II linear GSS 29-AUG-2000  
 typhimurium genomic clone 390-T3, genomic survey sequence.

ACCESSION AF075981  
 VERSION AF075981.1 GI:3320851  
 KEYWORDS GSS.  
 SOURCE Salmonella typhimurium  
 ORGANISM Salmonella typhimurium  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Salmonella.

REFERENCE 1 (bases 1 to 860)  
 AUTHORS Wong, R.M.-Y., Wong, K.K., Benson, N.R. and McClelland, M.  
 TITLE Sample sequencing of a Salmonella typhimurium LT2 lambda library:  
 comparison to the Escherichia coli K12 genome  
 JOURNAL FEMS Microbiol. Lett. 173 (2), 411-423 (1999)  
 MEDLINE 99243757  
 PubMed 10227170

COMMENT Contact: McClelland M  
 Molecular Biology  
 Sidney Kimmel Cancer Center  
 3099 Science Park Road, San Diego, CA 92121, USA  
 Email: mclelland@lifsci.sdsu.edu  
 Class: shotgun.

FEATURES  
 source  
 1. .860  
 /location="Qualifiers  
 /organism="Salmonella typhimurium"  
 /mol\_type="genomic DNA"  
 /strain="LT2"  
 /db\_xref="taxon:602"  
 /clone="390-T3"  
 /clone\_lib="Salmonella typhimurium LT2, Lambda DASH II"  
 /note="Vector: Lambda DASH II; sequenced using Li-Cor  
 sequencer"

ORIGIN  
 Query Match 3.0%; Score 72; DB 28; Length 860;  
 Best Local Similarity 54.7%; Pred. No. 9.9e-08;  
 Matches 220; Conservative 0; Mismatches 160; Indels 22; Gaps 3;

QY 1925 GCAGACCCATCGACGCGCATCTGCGCAGTGAAGTCTGCGGTTCTTCTGCGCATGTCA 1984  
 Db 400 CGCAGATCCAGCGAGCGCAATTTGATCAATTTTGAACCCGTCGCGCAGCGTTCA 341  
 QY 1985 TTGAGCTTGGCAGCATATCAATATTGTTTCAAGTCCCTCAGACAGGAGTTA 2044  
 Db 340 TTCAACCGGGAATTTTTCAGTTTCTGCGGGTAAATCTGTTTGAAGGAACGG 281  
 QY 2045 CCCTGGGTG---GCGCGGGAACCTCTGGAATTCATCGAGATATTGTCGCGTACGAGG 2101  
 Db 280 CTTCCGCTCGCGCGCGGGAATCTTCGGAATACCAATTCAAATTACGGTCACTGAGC-TG 222  
 QY 2102 CCCTGCGCAAGTGTGAGAAAGCAATGAGCCCAAGACCATTTGTCAGCTGACTGCAAC 2161  
 Db 221 CCGCGCCCGAGCGGAGAGGAAATCTACTACGAGCTTTTCTTCTGCAACAGCGCAGC 162  
 QY 2162 AAGTTCTTACCGTCATCGCCCGTTCTCTCCACCCCAACAAATTAATGATGAAATAGCTTGGC 2221  
 Db 161 AG-----CCCCGTCCTCTTACCCACGCTCAAAAGCGAAATATTAGGC 120  
 QY 2222 TGATGATCAGAGCGGAGCGCTCTCTCGCGATGAAGTCAAGCGCTCCCGTGTGAGC 2281  
 Db 119 TGATGATCAGAGCGGAGCGCTCGCGAGATCTCTCAGAAATATCGATAGCTTGTCTGCCAGA 60

QY 2282 TCTGACCGTAGAAGAAATACCCAGCTTAAGAGCGCTTTCCA 2323  
 Db 59 TCGCAGGATATTAGATCCCCACGATACACGCTTTGCCA 18

RESULT 4  
 CNS01MQH/c  
 LOCUS  
 DEFINITION CNS01MQH 830 bp DNA linear GSS 14-JUN-2001  
 Anopheles gambiae GSS T7 end of clone 22E24 of NotreDamel library  
 from strain PEST of Anopheles gambiae (African malaria mosquito),  
 genomic survey sequence.

ACCESSION ALI51258  
 VERSION ALI51258.1 GI:7011737  
 KEYWORDS GSS.  
 SOURCE Anopheles gambiae (African malaria mosquito)  
 ORGANISM Anopheles gambiae

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
 Anopheles.

REFERENCE 1 (bases 1 to 830)  
 Genoscope.  
 Direct Submission  
 Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : Segref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)

REFERENCE 2 (bases 1 to 830)  
 Authors Roth, C.W., Brey, P.T., Ke, Z., Collins, F.H. and Weissenbach, J.  
 Title Direct Submission  
 Submitted (16-FEB-2000) BEM, Institut Pasteur, 25, rue du Dr.  
 Roux, Paris 75015, France  
 Comment This clone is from an A. gambiae BAC library provided by F.H.  
 Collins and sequenced by Genoscope in collaboration with the  
 Laboratory of Biochem. and Biol. Molec. of Insects, Institut  
 Pasteur.

FEATURES  
 source  
 1. .830  
 /location="Qualifiers  
 /organism="Anopheles gambiae"  
 /mol\_type="genomic DNA"  
 /strain="PEST"  
 /db\_xref="taxon:7165"  
 /clone="22E24"  
 /clone\_lib="NotreDamel"  
 /note="end : T7"

ORIGIN  
 Query Match 2.5%; Score 60.4; DB 29; Length 830;  
 Best Local Similarity 50.7%; Pred. No. 0.00013;  
 Matches 142; Conservative 1; Mismatches 137; Indels 0; Gaps 0;

QY 1038 TTACAGGTCTGCTTTTGGGGCCAGTCTTTTACTGTCATCGACCGCAGATGTACTGG 1097  
 Db 597 KTCAGGGCTGCTTTAGTGTGATTTTATCTTCTCCCTCGCGCCAAAATGGTTTCG 538  
 QY 1098 TGATTAACAAGAAATTAAGCGCAAGACTCATTCGCGGTTCTTCGTCGTTAAATT 1157  
 Db 537 TGATGAATCAGGGCATCGCCGCCAGTACCACTGATGATTCGCTGCTATGCGGGTAA 478  
 QY 1158 CTGACGCTTTTGTTCATCGCGCCACCTTGGCGGTTGATCTTTTGTCCATGCGCGC 1217  
 Db 477 GTGATTTGCTGCTGATCTGTGCGCGGATTTTGGCGGCGCGCTGCTGATGCTGCTC 418  
 QY 1218 CGATCGTCTCGATATTATATGCGCTGGGGTGGCATCGCTTACCTTTATGTTTGGCTCA 1277  
 Db 417 CGTGGCTGCTGGCTTGGTGGTCCACCTGGGGCGGCTGGCTTCTGCTCTGTCGATTCG 358  
 QY 1278 TGGCAGCGAAGAGCGCCATGACAAACAGGTGGAGCGCC 1317  
 Db 357 GCGCGCTGAAACGCGGATGACAGTAACTCGAACTGGC 318

RESULT 5  
 BZ554294  
 LOCUS  
 BZ554294 1034 bp DNA linear GSS 17-DEC-2002

```

DEFINITION pac81-60_4617.xl pac81-60 Pseudomonas aeruginosa genomic clone
ACCESSION BZ554294
VERSION BZ554294.1 GI:27161466
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE 1 (bases 1 to 1034)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES             Location/Qualifiers
     source           1..1034
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="1-60"
                     /db_xref="taxon:287"
                     /clone="pac81-60_4617"
                     /clone_lib="pac81-60"
                     /note="clinical isolate 1-60 Whole genomic shotgun
                     library."
ORIGIN
Query Match      2.4%; Score 56.8; DB 28; Length 1034;
Best Local Similarity 50.4%; Pred. No. 0.0014;
Matches 139; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
QY 1022 ATCATGGAAATCTTCAATACAGTCTGCTTTGGGGCCAGTCTTTTACTGTCATCGGTTCTT 1141
DB 82 ATGTTTCAGAGCTATCTCAACGCATCTCTGGTGGCGGCTTGATATCGCCATTCGCG 141
QY 1082 CCGCAGATGCTACTGTGATTAACAAGAAATTAACGCGAAGACTCAATTCGGTGTCTT 1141
DB 142 GCGCAGAAAGCTTCTCTCCGCGAGAGCTCGCGCGGAGCATCACCTCTGGTAGCC 201
QY 1142 CTCGTGTGTTTAATTTCTACGCTCTTTTGTTCATCGCGCGGACCTTGGGCGTTGATCTT 1201
DB 202 GCGCTCTCGGTGTTCTCGACGCGGTGCTGCTCAGCTCGGCGGTGTTTCGCGCTGCGCAAG 261
QY 1202 TTGTCCATGCGCGCGGATCTGCTCGATATTAATCGTGGGTGGCATCGCTTACCTG 1261
DB 262 CTCGTGCTGGAAACCGACGCTGCTGGCCATCGCCGCTGGGGCGGAATCGGCTTCCTG 321
QY 1262 TTATGTTTGGCGTCAATGCGCAGCAAGACGCCATG 1297
DB 322 ACCGTGTAGGGCTCAAGGGCTGCTTCGCGGCTTG 357

RESULT 6
BZ560550/c
LOCUS BZ560550
DEFINITION pac82-164_2569.xl pac82-164 Pseudomonas aeruginosa genomic clone
ACCESSION BZ560550
VERSION BZ560550.1 GI:27178749
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE 1 (bases 1 to 897)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES             Location/Qualifiers
     source           1..897
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pac82-164_2569"
                     /clone_lib="pac82-164"
                     /note="clinical isolate 2-164 Whole genomic shotgun
                     library."
ORIGIN
Query Match      2.3%; Score 54.2; DB 28; Length 897;
Best Local Similarity 52.8%; Pred. No. 0.0067;
Matches 143; Conservative 0; Mismatches 120; Indels 8; Gaps 1;
QY 1455 TGATGGCAATCGTCTGACCTGGTTGAACCGAATCGGTATTTCGACGCGTTTGTGTTTA 1514
DB 552 TCGCGGCGACTGGGCTTCACTTGGCTTAACCCCACTCTATTCATACGCTACTTCTTAC 493
QY 1515 TCGCGGCGCTCGCGCGCAATACGCGCACACCGGACGCTGGAATTTTCGCGCTGGCGGT 1574
DB 492 GGTTCGCTCGCGCGCCAGCAAGNCGCG-----CCGCGCGCNTATGCTCGCGCGG 441
QY 1575 TCGCGCGAAGCTGATCTGCTTCGCTGCTGGTTCGCGCGCAGCAGCATTCGTACGCC 1634
DB 440 CCAAGCGCTCGCTGATGTGTTCTTCGCGCTCGCCCTCGCGCGGCGCATGCTGCCCCCT 381
QY 1635 CGCTGTCAGACCCCAAGGTGCGCGTGCATCAACCTCGTGCAGCTTGTGATGACCG 1694
DB 380 GGTTCGCGCGCGCGCGCACTGCGCGCTGCTCGACCTGATGCTGCGCGCCATGCTGCG 321
QY 1695 CATTCGCCATCAAACTGATGTTGATGGTTA 1725
DB 320 GCATGCGCGCGCAACTGCTGTTCCGCGGATA 290

RESULT 7
BZ568946/c
LOCUS BZ568946
DEFINITION pac82-164_8165.y2 pac82-164 Pseudomonas aeruginosa genomic clone
ACCESSION BZ568946
VERSION BZ568946.1 GI:27202770
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE 1 (bases 1 to 1620)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954

```

```

AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES             Location/Qualifiers
     source           1..897
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pac82-164_2569"
                     /clone_lib="pac82-164"
                     /note="clinical isolate 2-164 Whole genomic shotgun
                     library."
ORIGIN
Query Match      2.3%; Score 54.2; DB 28; Length 897;
Best Local Similarity 52.8%; Pred. No. 0.0067;
Matches 143; Conservative 0; Mismatches 120; Indels 8; Gaps 1;
QY 1455 TGATGGCAATCGTCTGACCTGGTTGAACCGAATCGGTATTTCGACGCGTTTGTGTTTA 1514
DB 552 TCGCGGCGACTGGGCTTCACTTGGCTTAACCCCACTCTATTCATACGCTACTTCTTAC 493
QY 1515 TCGCGGCGCTCGCGCGCAATACGCGCACACCGGACGCTGGAATTTTCGCGCTGGCGGT 1574
DB 492 GGTTCGCTCGCGCGCCAGCAAGNCGCG-----CCGCGCGCNTATGCTCGCGCGG 441
QY 1575 TCGCGCGAAGCTGATCTGCTTCGCTGCTGGTTCGCGCGCAGCAGCATTCGTACGCC 1634
DB 440 CCAAGCGCTCGCTGATGTGTTCTTCGCGCTCGCCCTCGCGCGGCGCATGCTGCCCCCT 381
QY 1635 CGCTGTCAGACCCCAAGGTGCGCGTGCATCAACCTCGTGCAGCTTGTGATGACCG 1694
DB 380 GGTTCGCGCGCGCGCGCACTGCGCGCTGCTCGACCTGATGCTGCGCGCCATGCTGCG 321
QY 1695 CATTCGCCATCAAACTGATGTTGATGGTTA 1725
DB 320 GCATGCGCGCGCAACTGCTGTTCCGCGGATA 290

RESULT 7
BZ568946/c
LOCUS BZ568946
DEFINITION pac82-164_8165.y2 pac82-164 Pseudomonas aeruginosa genomic clone
ACCESSION BZ568946
VERSION BZ568946.1 GI:27202770
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE 1 (bases 1 to 1620)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954

```



Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.  
Location/Qualifiers  
1. .1620

## FEATURES

source

/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="2-164"  
/db\_xref="taxon:287"  
/clone="pac82-164\_8165"  
/clone\_lib="pac82-164"  
/note="clinical isolate 2-164 Whole genomic shotgun library."

## ORIGIN

Query Match 2.2%; Score 51.4; DB 28; Length 1620;  
Best Local Similarity 54.3%; Pred. No. 0.052;  
Matches 100; Conservative 0; Mismatches 84; Indels 0; Gaps 0;  
QY 1542 ACACCGAGCGGTGATTTTCGCGCTGCGGCTTCGCGGCAAGCTGATCTGTTCCCGC 1601  
Db 618 AGCCCGCGNCGGCGATATCCCTCGCGCGCGGCGGCGCTCGATGATGTTCTTCG 559  
QY 1602 TGGTGGGTTTCGCGCGACGACATTTGTCACGCGCGCTGTCCAGCCCAAGTGTGCGCT 1661  
Db 558 CCCTCGCCCTCGCGCGGCGATGCTGCGCCCTCGCTGCGCGCGCCGCCACCTCGCGCC 499  
QY 1662 GATCAACGTCGTCGCGGATGATGATGACCGATGCGCATCAAACTGATGTTGATGG 1721  
Db 498 TGCTGACATGATGTCGCGCGCATGATGCTGGGCGATGCGCGCGCGCAACTGCTGTCGGG 439  
QY 1722 GTTA 1725  
Db 438 GATA 435

## RESULT 8

CN5017SY/c

LOCUS

DEFINITION

Drosophila melanogaster genome survey sequence SP6 end of BAC

BACN37L08 of DrosBAC library from Drosophila melanogaster (fruit

fly), genomic survey sequence.

ACCESSION

AL108460

VERSION

1

KEYWORDS

GSS.

SOURCE

Drosophila melanogaster (fruit fly)

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1101)

AUTHORS

Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a

collaboration with the European Drosophila Genome Project (EDGP) -

http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC

library (Dros BAC) was made by Alain Billaud at CNRS (Centre

d'Etude du Polymorphisme Humain) with funding provided by a MRC

project grant. The DNA was prepared from embryos by Alain Bucheton

and Genevieve Payan. It has been constructed in the vector

pBelOBAC11.

FEATURES

source

Location/Qualifiers

1. .1101

/organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/db\_xref="taxon:7227"

/clone="BACN37L08"

/clone\_lib="DrosBAC"

/plasmid="pBelOBAC11"

/note="end : SP6"

## ORIGIN

Query Match 2.1%; Score 48.8; DB 29; Length 1101;  
Best Local Similarity 14.5%; Pred. No. 0.21;  
Matches 47; Conservative 156; Mismatches 121; Indels 0; Gaps 0;  
QY 1492 GTATTTGACGCGCTTGTGTTTATCGGCGCGCTCGCGCCCAATACGCGACACCGAGC 1551  
Db 1079 KTTTBTBTSSBSSTSTSTBTBSBSSBSBBSBBSBSSBSSBSSBSSBSSBSSBSS 1020  
QY 1552 GTGATTTTCGCGCTGCGCGGTTCCGCGACGCTGATCTGTTCCCGCTGGTGGTTT 1611  
Db 1019 SSTSBTBTSTSSSTTTTTTTTTTTTBTBSBSSBSSBSSBSSBSSBSSBSSBSS 960  
QY 1612 CGGCGACGACGATTTGCAAGCCGCTGTCAGCCGCTCGAGCCCAAGCTGTCGCTGATCAACGT 1671  
Db 959 BSTSSASBSSSSSSBSSSTTTBTBSBBSSTSSSSSSSSBSSBSSBSSBSSBSSBT 900  
QY 1672 CGTCGCGGCGATTTGATGACCGCATTTGCCCATCAATCATGATGATGATGATGAT 1731  
Db 899 BTXSTSTSSST 840  
QY 1732 CGCGGTTTGGATCGGTGCGCTTCGCCCAATGTTGATCGCGCGCTGTCGGAATCTC 1791  
Db 839 TSGTBTBTBMSKBSSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 780  
QY 1792 ATCGATCGCTCAACTCGCGCTC 1815  
Db 779 GBCSTGCGCCCTCCCTCCTC 756

## RESULT 9

BZ556963/c

LOCUS

DEFINITION

pac81-60\_5975.x1 pac81-60 Pseudomonas aeruginosa genomic clone

pac81-60\_5975, genomic survey sequence.

ACCESSION

BZ556963

VERSION

BZ556963.1

KEYWORDS

GSS.

SOURCE

Pseudomonas aeruginosa

ORGANISM

Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

Pseudomonadaceae; Pseudomonas.

REFERENCE

1 (bases 1 to 833)

AUTHORS

Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,

Burns, J.L., Kaul, R. and Olsen, M.V.

TITLE

Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

COMMENT

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

Location/Qualifiers

1. .833

/organism="Pseudomonas aeruginosa"

/mol\_type="genomic DNA"

/strain="1-60"

/db\_xref="taxon:287"

/clone="pac81-60\_5975"

/clone\_lib="pac81-60"

/note="clinical isolate 1-60 Whole genomic shotgun library."

## ORIGIN

Query Match

Best Local Similarity

Matches 117; Conservative

2.0%; Score 48.6; DB 28;

Pred. No. 0.21;

Mismatches 114; Indels

0; Gaps

0;

QY 1067 TTAATGTCATCGGACCGGAGATGATCTGTTGATTAACAGAGATTAAGCGGAGCA 1126  
 Db 446 TCATCGGCATCGGCGGAGAGATGCTTCTGTCCTCGGAGAGCGCTCGCGCGAGCAT 387  
 QY 1127 CTCATTCGGGTCTTCTGCTGCTGTTAAATTTGATGACCTTTTGTTCATCGCGGAC 1186  
 Db 386 CACCTCTCGGTAGCGCGCTCTGCTGTTCTGCGACACGCTGCTGCTGCTGCTGCTG 327  
 QY 1187 TTGGGCGTGTGATCTTTTTCATGCGCGCGGATGCTGCTGCTGCTGCTGCTGCTG 1246  
 Db 326 TTGGGCGTGTGATCTTTTTCATGCGCGGATGCTGCTGCTGCTGCTGCTGCTG 267  
 QY 1247 GCGATCGCTTACTGTTATGTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1297  
 Db 266 GCGATCGCTTACTGTTATGTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 216

RESULT 10  
 CNS0091P 925 bp DNA linear GSS 03-JUN-1999  
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #  
 DEFINITION BAC1916 of RPI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
 ACCESSION AL053013.1 GI:4934461  
 VERSION AL053013.1  
 KEYWORDS GSS.  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 925)  
 AUTHORS Direct Submission  
 TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr  
 COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
 source  
 1. 925  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BACR1916"  
 /clone\_lib="RPI-98"  
 /note="end : TET3"

ORIGIN  
 Query Match 2.0%; Score 48.6; DB 29; Length 925;  
 Best Local Similarity 13.8%; Pred. No. 0.22; Indels 0; Gaps 0;  
 Matches 49; Conservative 157; Mismatches 149; Indels 0; Gaps 0;  
 QY 1340 CCAACGCTGCGGATGACAGCGCTTTGCGCGGTTGCGCGGCTGACCGGACAC 1399  
 Db 571 SCSCGSCGSCSSSCSCBCCGCCSSVCCSSSSSSKCSSTSBSGCCSSKSVCGTS 630  
 QY 1400 CGGTCGCGGTGAGTGAGCTGCTGATAGACGCGGTTTGGGTAAACCCATGTTGATG 1459  
 Db 631 CSBSSSCSSSSSTSSSTSSSTSSSSSSSSSSSTSSSTSSSSSSSSSSSSSSSSSS 690

QY 1460 GCAATCTGCTGACCTGTTGAACCCGAAATGCTATTGACGCGCTTGTGTTATCGGC 1519  
 Db 691 TGTGTSSTSSSSSTSSSSSVSSGSKSTSSBSBSSSSSSSSSTSSBSBCTSTSSSS 750  
 QY 1520 GCGCTCGCGCGCAATACCGGACACCGGACCGGTGATTTTCCGCTGCGCGCTGCG 1579  
 Db 751 SSVSSSTCSCTCCCTCCSVSSSTSSSSSTSSSSSTSSSSSTSSSSSTSSSSSTSS 810  
 QY 1580 GGAAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1639  
 Db 811 MCTCCSTYBCTSTSSGSSSGGKGVTKCGCGGSSSTNMBGTSSACSSSSSC 870  
 QY 1640 TCCAGCCCCAAGGTGCGCTGATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1694  
 Db 871 SSSSVSSSSKSSASSSSVSSSSSVSSSSSSKSSSSSVSSSSSVSSSSSVSSSSSV 925

RESULT 11  
 BX393687 852 bp mRNA linear EST 13-MAY-2003  
 LOCUS BX393687 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
 DEFINITION CDNA clone CS0DC001YG07 5-PRIME, mRNA sequence.  
 ACCESSION BX393687  
 VERSION BX393687.1 GI:30624032  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 852)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: segre@genoscope.cns.fr, Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : [liang@lifetech.com](mailto:liang@lifetech.com) URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Paraday Avenue Genoscope, sequence ID : CS0DC001AD04QF1.

FEATURES  
 source  
 1. 852  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DC001YG07"  
 /tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 2.0%; Score 47.6; DB 13; Length 852;  
 Best Local Similarity 17.2%; Pred. No. 0.4;  
 Matches 69; Conservative 152; Mismatches 178; Indels 3; Gaps 1;  
 QY 1038 TTACAGTCTGCTTTTGGGCGGCGAGTCTTTTCTGCTGCTGCTGCTGCTGCTGCTG 1097  
 Db 381 TTTCKGKTBKTKTKTKGKDKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 440  
 QY 1098 TGATTAACACAGGATTAAGCGGAGGACTCATTTGCTGCTGCTGCTGCTGCTGCTG 1157  
 Db 441 KKK 500  
 QY 1158 CTGACGCTCTTTTGTTCATCGCGGACCTTGGGCGGTGATCTTTGTCAGTCCGCGC 1217  
 Db 501 KDKKK 560  
 QY 1218 CGATCGTGTCTGATATTATGCGCTGGGCTGGCATCGCTTACCTGTTATGTTGCTG 1277



LOCUS BQ704105 644 bp mRNA linear EST 16-JUL-2002  
 DEFINITION 346110C04.y1.946 - tassell primordium prepared by Schmidt lab Zea  
 mays cDNA, mRNA sequence.  
 ACCESSION BQ704105  
 VERSION BQ704105.1 GI:21843524  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 644)  
 AUTHORS Walbot V.  
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
 UNIVERSITY  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 946110 row: C column: 04.  
 FEATURES  
 source  
 1. 644  
 Location/Qualifiers  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="OH43"  
 /db\_xref="taxon:4577"  
 /tissue\_type="tassels"  
 /dev\_stage="just after the transition from vegetative to  
 inflorescence development"  
 /lab\_host="XL0LX"  
 /clone\_lib="946 - tassell primordium prepared by Schmidt  
 lab"  
 /note="Organ: tassels; Vector: HybrizAP; Site 1: EcoRI;  
 Site 2: XhoI; George Chuck dissected immature tassels  
 between 1mm and 3mm. Sharon Stanfield prepared the cDNA  
 library in HybrizAP. Sample insert size range was 350 bp  
 to 3 Kb with a 1 Kb average."  
 ORIGIN  
 Query Match 1.9%; Score 44.4; DB 13; Length 644;  
 Best Local Similarity 46.0%; Pred. No. 2.5;  
 Matches 150; Conservative 0; Mismatches 176; Indels 0; Gaps 0;  
 QY 1364 TTGGCGGTTTCGGCGTGGCCACTGACACGCGCAACCGGTCGGGTGGAGTGGAGTGC 1423  
 DB 364 TTGGCTTGGACCGCGGAGAGTCCGAGATGGCGCCCTTGGAGAGCTTGGCGTTATGGTC 305  
 QY 1424 GATAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGTTGAAC 1483  
 DB 304 CTCCTCTAGTAGACGAGTGGAGCCGAGAGTGGAGTGGAGTGGAGTGGATC 245  
 QY 1484 CCGAATGCGTATTTCGACGCGTTTGTGTTATTCGCGCGCGTTCGCGCGCAATACGCGAC 1543  
 DB 244 TCGAAGCGCGCGTGGCTTCTCGAAGCTGTAGGAGCGCAGGAGTGGCGGATGAGCCCC 185  
 QY 1544 ACCGACGCTGGATTTTCGCGTGGCGGTTTCGCGCGAGCTGATCTGCTCCCGCTG 1603  
 DB 184 TTCGGGAGGCGCGTACTTGGGAGTAGTCTTTCGCGCGCGTGGAGCGATGCTGCTGAC 125  
 QY 1604 GTGGGTTTCGGCGCAGCAGCATTTGTCAACCGCGTGTTCAGCCCCCAAGGTGTGGCGTGG 1663  
 DB 124 GCCGTGACCGCTCGA 65  
 QY 1664 ATCAACGCTCGTGTGGCAGTTGTGAT 1689  
 DB 64 GACATTGCCGTTGGGATATTGGAT 39

Search completed: March 16, 2004, 02:52:12

RESULT 14  
 CAB29022/c  
 LOCUS CAB29022 620 bp mRNA linear EST 11-DEC-2002  
 DEFINITION 1114036D07.y2.1114 - Unigene IV from Maize Genome Project Zea mays  
 cDNA, mRNA sequence.  
 ACCESSION CAB29022  
 VERSION CAB29022.1 GI:26457439  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 620)  
 AUTHORS Walbot V.  
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
 UNIVERSITY  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 1114036 row: D column: 07.  
 FEATURES  
 source  
 1. 620  
 Location/Qualifiers  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4577"  
 /clone\_lib="1114 - Unigene IV from Maize Genome Project"  
 /note="This library represents the unique genes found in  
 the fourth round of EST sequencing at Stanford University  
 for the maize genome project. Sequences are present from  
 libraries 1091 and 3524. Contigs were assembled using  
 ZmBAssembler and 2 representatives from each contig were  
 selected for the Unigene set. All singlets were also  
 selected."  
 ORIGIN  
 Query Match 1.9%; Score 44.4; DB 14; Length 620;  
 Best Local Similarity 46.0%; Pred. No. 2.5;  
 Matches 150; Conservative 0; Mismatches 176; Indels 0; Gaps 0;  
 QY 1364 TTGGCGGTTTCGGCGTGGCCACTGACACGCGCAACCGGTCGGGTGGAGTGGAGTGC 1423  
 DB 357 TTGGCTTGGACCGCGGAGAGTCCGAGATGGCGCCCTTGGAGAGCTTGGCGTTATGGTC 298  
 QY 1424 GATAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGTTGAAC 1483  
 DB 297 CTCCTCTAGTAGACGAGTGGAGCCGAGAGTGGAGTGGAGTGGAGTGGATC 238  
 QY 1484 CCGAATGCGTATTTCGACGCGTTTGTGTTATTCGCGCGCGTTCGCGCGCAATACGCGAC 1543  
 DB 237 TCGAAGCGCGCGTGGCTTCTCGAAGCTGTAGGAGCGCAGGAGTGGCGGATGAGCCCC 178  
 QY 1544 ACCGACGCTGGATTTTCGCGTGGCGGTTTCGCGCGAGCTGATCTGCTCCCGCTG 1603  
 DB 177 TTCGGGAGGCGGCTACTTGGGAGTAGTCTTTCGCGCGCGCTGAGGAGCTGGCTGAC 118  
 QY 1604 GTGGGTTTCGGCGCAGCAGCATTTGTCAACCGCGCTTCAGCCCCCAAGGTGTGGCGTGG 1663  
 DB 117 GCCGTGACCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGA 58  
 QY 1664 ATCAACGCTCGTGTGGCAGTTGTGAT 1689  
 DB 57 GACATTGCCGTTGGGATATTGGAT 32

RESULT 15  
 BQ704105/c

Thu Mar 18 12:31:00 2004

us-09-105-117k-1.rst

Page 9

Job time : 6533.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:15:20 ; Search time 2846.7 Seconds  
(without alignments)  
10825.501 Million cell updates/sec

Title: US-09-105-117K-1 COPY 1016 1726

Accession: **CC 03 163 117A 1\_2010\_1720**  
 Perfect score: **711**  
 Sequence: **1 atggtcaatcatggaatctt.....aactgatattgatgggttag 711**

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:★

1: gb\_ba.\*  
2: gb\_ba.\*  
3: gb\_int.\*  
4: gb\_int.\*  
5: gb\_ov.\*  
6: gb\_ov.\*  
7: gb\_pa.\*  
8: gb\_pa.\*  
9: gb\_ph.\*  
10: gb\_ph.\*  
11: gb\_ro.\*  
12: gb\_ro.\*  
13: gb\_scs.\*  
14: gb\_scs.\*  
15: gb\_sy.\*  
16: gb\_sy.\*  
17: gb\_un.\*  
18: gb\_un.\*  
19: gb\_vl.\*  
20: gb\_vl.\*  
21: gb\_vl.\*  
22: gb\_vl.\*  
23: gb\_vl.\*  
24: gb\_vl.\*  
25: gb\_vl.\*  
26: gb\_vl.\*  
27: gb\_vl.\*  
28: gb\_vl.\*  
29: gb\_vl.\*  
30: gb\_vl.\*  
31: gb\_vl.\*  
32: gb\_vl.\*  
33: gb\_vl.\*  
34: gb\_vl.\*  
35: gb\_vl.\*  
36: gb\_vl.\*  
37: gb\_vl.\*  
38: gb\_vl.\*  
39: gb\_vl.\*  
40: gb\_vl.\*  
41: gb\_vl.\*

pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	711	100.0	711	6	AX643028	Sequence
2	711	100.0	822	6	AX063771	Sequence
3	711	100.0	822	6	AX244059	Sequence
4	711	100.0	2374	1	CXYLSSBG	C, glutamic
5	711	100.0	2374	6	A93933	Sequence 2
6	711	100.0	33150	1	AP005277	Corynebac
7	711	100.0	34980	6	AX127147	Sequence
8	708	99.6	708	6	AX123539	Sequence
9	708	99.6	708	6	BD165656	Novel pol
10	696.8	98.0	712	6	AX643030	Sequence
11	317.6	44.7	1568	6	E54483	Heat-resist
12	317.6	44.7	1771	1	AB083133	Corynebac
13	317.6	44.7	308650	1	AP005218	Corynebac
14	136.6	19.2	349535	1	BX248357	Corynebac
15	84.2	11.8	190050	1	AL646059	Ralstonia
16	65.8	9.3	300861	1	AE016777	Pseudomon
17	65.6	9.2	303550	1	SC0939131	Streptomy
18	63.4	8.9	10542	1	AE004852	Pseudomon
19	63.2	8.9	10024	1	AE015725	Shewanell
20	62.8	8.8	24336	1	AE008841	Salmonell
21	62.8	8.8	298300	1	AP005025	Streptomy
22	62.2	8.7	311600	1	AE016871	Pseudomon
23	62	8.7	10313	1	AE015304	Shigella
24	62	8.7	10701	1	AE005522	Escherich
25	62	8.7	266658	1	AP002563	Escherich
26	62	8.7	292906	1	AE016988	Shigella
27	61.6	8.7	4033	1	AS065741	Aeromonas s
28	61.2	8.6	230050	1	AL627277	Salmonell
29	61.2	8.6	301574	1	AE016844	Salmonell
30	60.4	8.5	636	6	E49392	Process for
31	60.4	8.5	636	6	E49392	Process for
32	60.4	8.5	636	6	AX030085	Sequence
33	60.4	8.5	8029	1	ECFDAPKG	Escherichia
34	60.4	8.5	10362	1	AE000375	Escherichia
35	60.4	8.5	141744	1	ECT28377	Gallus
36	60.4	8.5	241182	2	AC145927	Gallus
37	59.6	8.4	213050	1	AL646067	Ralstonia
38	58.8	8.4	10747	1	AE012410	Xanthomon
39	58	8.3	303121	1	AE016766	Escherich
40	57	8.0	14844	1	MBU34849	Mycobacteri
41	57	8.0	14869	1	AE007056	Mycobacter
42	57	8.0	15239	6	I86263	Sequence 17
43	57	8.0	306050	1	BX248341	Mycobacte
44	56.6	8.0	346186	1	EX842578	Sequence
45	56.6	8.0	642	6	AP387945	Sequence
					AP005341	Vibri

## ALIGNMENTS

RESULT 1					
AX643028					
LOCUS		711 bp	DNA	linear	PAT 24-FEB-2003
DEFINITION	Sequence 7 from Patent EP1266956.				
ACCESSION	AX643028				
VERSION	AX643028.1	GI:28550158			
KEYWORDS	Corynebacterium glutamicum				
SOURCE	Corynebacterium glutamicum				
ORGANISM	Bacteria; Actinobacteria; Actinomycetales;				
	Corynebacterineae; Corynebacteriaceae; Corynebacterium.				
REFERENCE	1				
AUTHORS	Gunji, Y. and Yasueda, H.				
TITLE	Method for producing L-lysine or L-arginine by using methanol assimilating bacterium				

JOURNAL Patent: EP 1266966-A 7 18-DEC-2002;  
 Ajinomoto Co., Inc. (JP)  
 FEATURES Location/Qualifiers  
 source 1..711  
 /organism="Corynebacterium glutamicum"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1718"  
 CDS 1..711  
 /notes="unnamed protein product"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAD67811.1"  
 /db\_xref="GI:28550159"  
 /db\_xref="REMBL:CAD67811"  
 /translation="MVEIFITGLLGSALLSIGQNVVIVKQIKRGLIALLV  
 CLISDFVFIAGTGLVLSNAAPVILDMRWGGIAYLLWFAVMAKADMTNKVEAPQ  
 IIEETFPDPTPLGSAVATDRNRVREVSVDKRVKPMALVITLWPNAY  
 LDFAVFGVGAQVGDTRWIFAAAGFAASLWPLVFGFAAALSRLSPSSPKWRWIN  
 VVAVVMTALAIXLMNG"  
 ORIGIN  
 Query Match 100.0%; Score 711; DB 6; Length 711;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-169;  
 Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGTGATCATGAATCTTCATTACAGGTCTGCTTTGGGGCCAGTCTTTTACTGTCC 60  
 DB 1 ATGGTGATCATGAATCTTCATTACAGGTCTGCTTTGGGGCCAGTCTTTTACTGTCC 60  
 QY 61 ATCGGACCGCAGAACTACTGTGATTAACAAGGAATTAAGCGCGAAGGACTCATTTGCG 120  
 DB 61 ATCGGACCGCAGAACTACTGTGATTAACAAGGAATTAAGCGCGAAGGACTCATTTGCG 120  
 QY 121 GTTCTTCTCGTGTGTTAAATTTCTGAGCTCTTTTGTTCATCGCCGGCACCTTGGGCGTT 180  
 DB 121 GTTCTTCTCGTGTGTTAAATTTCTGAGCTCTTTTGTTCATCGCCGGCACCTTGGGCGTT 180  
 QY 181 GATCTTTTGTCCAAATCGCGCGCATCGTCTCGATATTATGCGCTGGGCTGGCATCGCT 240  
 DB 181 GATCTTTTGTCCAAATCGCGCGCATCGTCTCGATATTATGCGCTGGGCTGGCATCGCT 240  
 QY 241 TACCTGTATGTTTTCGCTCATGCGCAGGAAAGCGCATGACAAACAAGTGAAGCG 300  
 DB 241 TACCTGTATGTTTTCGCTCATGCGCAGGAAAGCGCATGACAAACAAGTGAAGCG 300  
 QY 301 CCACAGATCATTAAGAAGAAACAGAACCGCTGCGGATGACGCGTGGGCGGTTTCG 360  
 DB 301 CCACAGATCATTAAGAAGAAACAGAACCGCTGCGGATGACGCGTGGGCGGTTTCG 360  
 QY 361 GCGGTGCGCACTGACGCGGCAACCGGCTGCGGCTGGAGTGAAGTGAAGCAGCGG 420  
 DB 361 GCGGTGCGCACTGACGCGGCAACCGGCTGCGGCTGGAGTGAAGTGAAGCAGCGG 420  
 QY 421 GTTTGGGTAAAGCCAGTGTGATGAGCAATCGTGTGATGAGTGAAGTGAAGTGAAG 480  
 DB 421 GTTTGGGTAAAGCCAGTGTGATGAGCAATCGTGTGATGAGTGAAGTGAAGTGAAG 480  
 QY 481 TTGGAGCGGTTTGTGTTATCGCGCGCTGCGGCGCAATACGCGGACACCGGAGTGG 540  
 DB 481 TTGGAGCGGTTTGTGTTATCGCGCGCTGCGGCGCAATACGCGGACACCGGAGTGG 540  
 QY 541 ATTTTCGCGCGTTCGCGCGTTCGCGCAAGCTGATCTGTTTCCGCTGTTGGGTTTCGCG 600  
 DB 541 ATTTTCGCGCGTTCGCGCGTTCGCGCAAGCTGATCTGTTTCCGCTGTTGGGTTTCGCG 600  
 QY 601 GCAGCAGATGTTGACGCGCGCTGTGACGCGGCGGATGAGTGTGCGCTGGATCAAGTCTGC 660  
 DB 601 GCAGCAGATGTTGACGCGCGCTGTGACGCGGCGGATGAGTGTGCGCTGGATCAAGTCTGC 660  
 QY 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 711  
 DB 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 711

RESULT 2  
 LOCUS AX063771 822 bp DNA linear PAT 24-JAN-2001  
 DEFINITION Sequence 53 from Patent WO0100843.  
 ACCESSION AX063771  
 VERSION AX063771.1 GI:12541483  
 KEYWORDS  
 SOURCE Corynebacterium glutamicum  
 ORGANISM Corynebacterium glutamicum  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 REFERENCE 1  
 AUTHORS Pompejus M., Kroeger B., Schroeder H., Zelder O. and Haberhauer G.  
 TITLE corynebacterium glutamicum genes encoding metabolic pathway  
 proteins  
 JOURNAL Patent: WO 0100843-A 53 04-JAN-2001;  
 BASF AKTIENGESELLSCHAFT (DE)  
 FEATURES Location/Qualifiers  
 source 1..822  
 /organism="Corynebacterium glutamicum"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1718"  
 101..802  
 /notes="unnamed protein product; RXA01394"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAC25127.1"  
 /db\_xref="GI:12541484"  
 /db\_xref="REMBL:CAC25127"  
 /translation="MEIFITGLLGSALLSIGQNVVIVKQIKRGLIALLVCL  
 SDVPLTAGTGLVLSNAAPVILDMRWGGIAYLLWFAVMAKADMTNKVEAPQ  
 ETEFTFPDPTPLGSAVATDRNRVREVSVDKRVKPMALVITLWPNAYLDA  
 FVETGVAQVGDTRWIFAAAGFAASLWPLVFGFAAALSRLSPSSPKWRWINVV  
 AVVMTALAIXLMNG"  
 ORIGIN  
 Query Match 100.0%; Score 711; DB 6; Length 822;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-169;  
 Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGTGATCATGAATCTTCATTACAGGTCTGCTTTGGGGCCAGTCTTTTACTGTCC 60  
 DB 92 ATGGTGATCATGAATCTTCATTACAGGTCTGCTTTGGGGCCAGTCTTTTACTGTCC 151  
 QY 61 ATCGGACCGCAGAACTACTGTGATTAACAAGGAATTAAGCGCGAAGGACTCATTTGCG 120  
 DB 152 ATCGGACCGCAGAACTACTGTGATTAACAAGGAATTAAGCGCGAAGGACTCATTTGCG 211  
 QY 121 GTTCTTCTCGTGTGTTAAATTTCTGAGCTCTTTTGTTCATCGCGGCGACCTTGGGCGTT 180  
 DB 212 GTTCTTCTCGTGTGTTAAATTTCTGAGCTCTTTTGTTCATCGCGGCGACCTTGGGCGTT 271  
 QY 181 GATCTTTTGTCCAAATCGCGCGCATCGTCTCGATATTATGCGCTGGGCTGGCATCGCT 240  
 DB 272 GATCTTTTGTCCAAATCGCGCGCATCGTCTCGATATTATGCGCTGGGCTGGCATCGCT 331  
 QY 241 TACCTGTATGTTTTCGCTCATGCGCAGGAAAGCGCATGACAAACAAGTGAAGCG 300  
 DB 332 TACCTGTATGTTTTCGCTCATGCGCAGGAAAGCGCATGACAAACAAGTGAAGCG 391  
 QY 301 CCACAGATCATTAAGAAGAAACAGAACCGCTGCGGATGACGCGTGGGCGGTTTCG 360  
 DB 392 CCACAGATCATTAAGAAGAAACAGAACCGCTGCGGATGACGCGTGGGCGGTTTCG 451  
 QY 361 GCGGTGCGCACTGACGCGGCAACCGGCTGCGGCTGGAGTGAAGTGAAGCAGCGG 420  
 DB 452 GCGGTGCGCACTGACGCGGCAACCGGCTGCGGCTGGAGTGAAGTGAAGCAGCGG 511  
 QY 421 GTTTGGGTAAAGCCAGTGTGATGAGCAATCGTGTGATGAGTGAAGTGAAGTGAAG 480  
 DB 512 GTTTGGGTAAAGCCAGTGTGATGAGCAATCGTGTGATGAGTGAAGTGAAGTGAAG 571



```

QY 481 TTGGAGCGCTTTGTTTATCGCGCGCTCGCGCGCAATACGGCGACACCGGACGGTGG 540
Db 572 TTGGAGCGCTTTGTTTATCGCGCGCTCGCGCGCAATACGGCGACACCGGACGGTGG 631
QY 541 ATTTTCGCGCTGGCGGCTTCGGCGGAGCTGATCTGGTTCCCGCTGGTGGGTTTCGGC 600
Db 632 ATTTTCGCGCTGGCGGCTTCGGCGGAGCTGATCTGGTTCCCGCTGGTGGGTTTCGGC 691
QY 601 GCAGCAGCATTTGATGACCGCGCTGTCAGCGCCCAAGGCTGGCGCTGGATCAACGTCGTC 660
Db 692 GCAGCAGCATTTGATGACCGCGCTGTCAGCGCCCAAGGCTGGCGCTGGATCAACGTCGTC 751
QY 661 GTGCGAGTTGTTGATGACCGCATTTGCGCATCAAACTGATGTTGATGGGTTAG 711
Db 752 GTGCGAGTTGTTGATGACCGCATTTGCGCATCAAACTGATGTTGATGGGTTAG 802

RESULT 3
AX244059
LOCUS 822 bp DNA linear PAT 28-SEP-2001
DEFINITION Sequence 51 from Patent WO0166573.
ACCESSION AX244059
VERSION AX244059.1 GI:15859123
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS Pompeius, M., Kroeger, B., Schroeder, H., Zelder, O., Haberbauer, G.,
Kim, J.W., Lee, H.S. and Hwang, B.J.
TITLE Corynebacterium glutamicum genes encoding metabolic pathway
JOURNAL Patent: WO 0166573-A 51 13-SEP-2001;
BASIS AKTIENGESCHSCHAFT (DE)
FEATURES
source
1..822
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101..802
/note="unnamed protein product; RXA01394"
/codon_start=1
/transl_table=11
/protein_id="CAC88508.1"
/db_xref="GI:15859124"
/db_xref="REFSEQ:CAC88508"
/translation="MEIFITGLLGLGALLSTGPNVLIVKQIKRGLIIVLVCLL
SDVFLIAGTGLVDLLSNAPIVLDIMRWGGIAYLLWFAVNAKADMTNKPQIIE
ETEPTVDDPLGGSAAVATDRNRVVEVDKQVVKPKMLAIVLWLNPNAYLDA
FVFIGGAGYQGTGRWIPFAGAFASLIWFLVGFAGAAALSRPLSPKVMRWINVVV
AVVMTALAIKMLMG"

ORIGIN
Query Match 100.0%; Score 711; DB 6; Length 822;
Best Local Similarity 100.0%; Pred. No. 2.7e-169;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGATCATCGAATCTTATTACAGGTCTGCTTTGGGGCGCAGCTTTTACTGTCC 60
Db 92 ATGGTGATCATCGAATCTTATTACAGGTCTGCTTTGGGGCGCAGCTTTTACTGTCC 151
QY 61 ATCGAGCGCAGATGTACTGGTATTAAACAAGGAATTAAGCGCGAAGGACTCATTTGG 120
Db 152 ATCGAGCGCAGATGTACTGGTATTAAACAAGGAATTAAGCGCGAAGGACTCATTTGG 211
QY 121 GTTCTCTCGTGTGTTAAATTCAGACGCTTTTGTTCATCGCGGACCTTGGGGGTT 180
Db 212 GTTCTCTCGTGTGTTAAATTCAGACGCTTTTGTTCATCGCGGACCTTGGGGGTT 271
QY 181 GATCTTTTCTCAATGCGCGCGATCGTCTCGATATTATCGCTGGGTGGCATCGCT 240
Db 272 GATCTTTTCTCAATGCGCGCGATCGTCTCGATATTATCGCTGGGTGGCATCGCT 331

```

```

QY 241 TACCTGTTATGTTTGGTTCATCGGAGCGAAAGACGCGCATGACAAACAGGTGGAAGCG 300
Db 332 TACCTGTTATGTTTGGTTCATCGGAGCGAAAGACGCGCATGACAAACAGGTGGAAGCG 391
QY 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGGGCGGTTTCG 360
Db 392 CCACAGATCATTTGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGGGCGGTTTCG 451
QY 361 CGGTTGGCGCTGATGACACCGCGACCGGGTGGGGTGGAGGTGAGGTGAGGTGAGGTGAGGTGAG 420
Db 452 CGGTTGGCGCTGATGACACCGCGACCGGGTGGGGTGGAGGTGAGGTGAGGTGAGGTGAGGTGAG 511
QY 421 GTTTGGGTTAAAGCCCATGTTGATGGCAATCTGCTGACCTGGTTGAACCCGCAATGCGGTAT 480
Db 512 GTTTGGGTTAAAGCCCATGTTGATGGCAATCTGCTGACCTGGTTGAACCCGCAATGCGGTAT 571
QY 481 TTGACGCGGTTTGTGTTATTCGCGCGCTGCGCGCGCAATACCGCGACACCGGACGGTGG 540
Db 572 TTGACGCGGTTTGTGTTATTCGCGCGCTGCGCGCGCAATACCGCGACACCGGACGGTGG 631
QY 541 ATTTTCGCGCTGGCGGCTTCGGCGGAGCTGATCTGTTCCGCGTGGTGGGTTTCGGC 600
Db 632 ATTTTCGCGCTGGCGGCTTCGGCGGAGCTGATCTGTTCCGCGTGGTGGGTTTCGGC 691
QY 601 GCAGCAGCATTTGTCAGCGCGCTGTCAGCGCCCAAGGCTGGCGCTGGATCAACGTCGTC 660
Db 692 GCAGCAGCATTTGTCAGCGCGCTGTCAGCGCCCAAGGCTGGCGCTGGATCAACGTCGTC 751
QY 661 GTGCGAGTTGTTGATGACCGCATTTGGCGCATCAAACTGATGTTGATGGGTTAG 711
Db 752 GTGCGAGTTGTTGATGACCGCATTTGGCGCATCAAACTGATGTTGATGGGTTAG 802

RESULT 4
CGLYSEG
LOCUS 2374 bp DNA linear BCT 19-MAR-2001
DEFINITION C-glutamicum lyse and lyse genes.
ACCESSION X96471
VERSION X96471.1 GI:1729753
KEYWORDS lyse gene; lyse gene; lysine export regulator protein; lysine
exporter protein; lysine governor.
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS Vrljic, M., Sahm, H. and Eggeling, L.
TITLE A new type of transporter with a new type of cellular function:
L-lysine export from Corynebacterium glutamicum
JOURNAL Mol. Microbiol. 22 (5), 815-826 (1996)
MEDLINE 97126810
PUBMED 8971704
REFERENCE
AUTHORS Vrljic, M.M.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-1996) M.M. Vrljic, Institut fuer Biotechnologie
1, Forschungszentrum Juelich, Postfach 1913, D-52425 Juelich, FRG
FEATURES
source
1..2374
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/strain="R127"
/db_xref="taxon:1718"
complement (82..954)
/gene="lyse"
complement (82..954)
/note="Lysine governor"
/codon_start=1
/transl_table=11
/product="Lysine export regulator protein"
/protein_id="CAA65323.1"

```

```

/db xref="GI:1729754"
/db xref="GOA:P94632"
/db xref="SWISS-PROT:P94632"
/translation="WNPIQLDILSLIIDESEPCASIALSISPSAVSORVKALEHHV
RVLSVTPAKATAGEVVLQARQWLLQAEIKAQISGRLEAPIPLIATINADSJSTW
FPVFNEVASGGATLRLDEADHTLSLLRRGDIAGVATREANPVAGCBVBLGTMR
HLAIAIPLSDAYVMDGKLDWAMPVLPFGKQVLDQRDLDRGVDPVRRRVSIVPS
AEGFGAIRRGGLGMLLPETQAPMLKAGVILLIDETPIDTPMYQWRLESRLARL
TDAVDAATIEGLRP"
gene
1025..1726
/ene="lyse"
1025..1726
/ene="lyse"
/codon_start=1
/transl_table=11
/product="Lysine exporter protein"
/protein_id="CAA65324.2"
/db xref="GI:13397387"
/db xref="GOA:P94633"
/db xref="SWISS-PROT:P94633"
/translation="WEIFITGLLGLASLLSIGPQNVLVIKQIKREGLIALLVCL
SDVLFIACTGLVLLSNAAPIVLDIMRWGGIAYLLWFAVMAAKDAMTKVBAQIIE
ETPTPTDPLDGSASVATCTNEVRVSVSDKQVWVWKMALVLTWLNPNAYLDA
FVIGVGAGYQDGRWIPFAGAPASLIWFLVFGGAALSRPLSSPKWRNVVW
AVWTAIAIKMLMG"
gene
complement(1723..2352)
/ene="orf3"
complement(1723..2352)
/ene="orf3"
/codon_start=1
/transl_table=11
/protein_id="CAA65325.1"
/db xref="GI:1729756"
/db xref="SPTREMBL:P94634"
/translation="MYALRDIVASGKALVGISSVGPETLTAABEFMAEBCPULLHQ
PSYIINRWEPDGDENLLQSAANGLGVIAISPLAQGLLTDKYLQDIPESRASQ
GKSLSGMLNVNNDIMKLNLDIAQERGSLAQWALVDRQGGYGIADIVTSALIGA
SSVEQLDNLSDLSNLEFSDAELEAIDEISHDAGINIWAKATDSKTREN"
ORIGIN
Query Match 100.0%; Score 711; DB 1; Length 2374;
Best Local Similarity 100.0%; Pred. No. 3.1e-169;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGTGATCATGGAATCTTCAATACAGGCTCTCTTTGGGGCCAGTCTTTACTGTCC 60
DB 1016 ATGGTGATCATGGAATCTTCAATACAGGCTCTCTTTGGGGCCAGTCTTTACTGTCC 1075
QY 61 ATCGACCGCAGATGCTGTGATTAACAGGAATTAAGCGCGAAGACTCATTTGG 120
DB 1076 ATCGACCGCAGATGCTGTGATTAACAGGAATTAAGCGCGAAGACTCATTTGG 1135
QY 121 GTTCTTCTCGTGTGTTTAAATTTCTGAGCTCTTTTTCATCGCCGACCTTGGCGCTT 180
DB 1136 GTTCTTCTCGTGTGTTTAAATTTCTGAGCTCTTTTTCATCGCCGACCTTGGCGCTT 1195
QY 181 GATCTTTTGTCCATGATCGCGCGGATGCTGATTAATGATGCTGCGGTGGCATGCTT 240
DB 1196 GATCTTTTGTCCATGATCGCGCGGATGCTGATTAATGATGCTGCGGTGGCATGCTT 1255
QY 241 TACCTGTTATGTTTGGTCCGCTCATGCGAGCGAAAGCGCCATGACAAAAGGTGGAAGCG 300
DB 1256 TACCTGTTATGTTTGGTCCGCTCATGCGAGCGAAAGCGCCATGACAAAAGGTGGAAGCG 1315
QY 301 CCACAGATCATTTAGAAACAGAACACGCGGTGCGGTGAGGTGAGCTGATAGCAGCG 420
DB 1316 CCACAGATCATTTAGAAACAGAACACGCGGTGCGGTGAGGTGAGCTGATAGCAGCG 1375
QY 361 GCGGTGGCCACTGACACGCGCAACCGGGTGGAGGTGAGCTGATAGCAGCG 420
DB 1376 GCGGTGGCCACTGACACGCGCAACCGGGTGGAGGTGAGCTGATAGCAGCG 1435
QY 421 GTTGGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGTTGTAACCGCATCGGTAT 480

```

```

DB 1436 GTTGGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGTTGAACCGCAATCGGTAT 1495
QY 481 TTGGACGGTTCGTGTTATCGGCGCGTTCGGCGCAATACGGCGACACCGGCGGTGG 540
DB 1496 TTGGACGGTTCGTGTTATCGGCGCGTTCGGCGCAATACGGCGACACCGGCGGTGG 1555
QY 541 ATTTTCGCGCTGCGCGCTTCGGCGCAAGCTGATCTGTTCCGCTGCGGTGCGGTTCGCG 600
DB 1556 ATTTTCGCGCTGCGCGCTTCGGCGCAAGCTGATCTGTTCCGCTGCGGTGCGGTTCGCG 1615
QY 601 GCAGCAGCATTTGACCGCGCTTCAGCGCGCTTCAGCGCGCGGCTGATGATCAACGTCGTC 660
DB 1616 GCAGCAGCATTTGACCGCGCTTCAGCGCGCGGCTTCAGCGCGCGGCTGATGATCAACGTCGTC 1675
QY 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAG 711
DB 1676 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAG 1726

RESULT 5
A93933
LOCUS 2374 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 2 from Patent WO9723597.
ACCESSION A93933
VERSION A93933.1 GI:6742037
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1. Ibañez J. and Eggeling L.
AUTHORS Vrijl J. M. and Eggeling L.
TITLE PROCESS FOR THE MICROBIAL PRODUCTION OF AMINO ACIDS BY BOOSTED
JOURNAL ACTIVITY OF EXPORT CARRIERS
Patent: WO 9723597-A 2 03-JUL-1997;
KERNFORSCHUNGSANLAGE JUELICH (DE); VRLIJC MARINA (DE)
FEATURES
source
1..2374
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

ORIGIN
Query Match 100.0%; Score 711; DB 6; Length 2374;
Best Local Similarity 100.0%; Pred. No. 3.1e-169;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGTGATCATGGAATCTTCAATACAGGCTCTCTTTGGGGCCAGTCTTTACTGTCC 60
DB 1016 ATGGTGATCATGGAATCTTCAATACAGGCTCTCTTTGGGGCCAGTCTTTACTGTCC 1075
QY 61 ATCGACCGCAGATGCTGTGATTAACAGGAATTAAGCGCGAAGACTCATTTGG 120
DB 1076 ATCGACCGCAGATGCTGTGATTAACAGGAATTAAGCGCGAAGACTCATTTGG 1135
QY 121 GTTCTTCTCGTGTGTTTAAATTTCTGAGCTCTTTTTCATCGCGGACCTTGGCGCTT 180
DB 1136 GTTCTTCTCGTGTGTTTAAATTTCTGAGCTCTTTTTCATCGCCGACCTTGGCGCTT 1195
QY 181 GATCTTTTGTCCATGATCGCGCGGATGCTGATTAATGATGCTGCGGTGGCATGCTT 240
DB 1196 GATCTTTTGTCCATGATCGCGCGGATGCTGATTAATGATGCTGCGGTGGCATGCTT 1255
QY 241 TACCTGTTATGTTTGGTCCGCTCATGCGAGCGAAAGCGCCATGACAAAAGGTGGAAGCG 300
DB 1256 TACCTGTTATGTTTGGTCCGCTCATGCGAGCGAAAGCGCCATGACAAAAGGTGGAAGCG 1315
QY 301 CCACAGATCATTTAGAAACAGAACACGCGGTGCGGTGAGGTGAGCTGATAGCAGCG 420
DB 1316 CCACAGATCATTTAGAAACAGAACACGCGGTGCGGTGAGGTGAGCTGATAGCAGCG 1375
QY 361 GCGGTGGCCACTGACACGCGCAACCGGGTGGAGGTGAGCTGATAGCAGCG 420
DB 1376 GCGGTGGCCACTGACACGCGCAACCGGGTGGAGGTGAGCTGATAGCAGCG 1435
QY 421 GTTGGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGTTGTAACCGCATCGGTAT 480

```

Db 1376 GCCTGGCCACTGACGCGCAACCGCGTGGCGGTGAGGTGAGCTGATAGCAGCGG 1435  
QY 421 GTTGGGTAAAGCCCATGTTGATGCGCAATCGTCTGACCTGCTGACCCGATCGGTAT 480  
Db 1436 GTTGGGTAAAGCCCATGTTGATGCGCAATCGTCTGACCTGCTGACCCGATCGGTAT 1495  
QY 481 TTGGACGCGTGTGTTGTTATCGCGCGCGCTCGCGCGCAATACCGCGACACCGACCGGTGG 540  
Db 1496 TTGGACGCGTGTGTTGTTATCGCGCGCGCTCGCGCGCAATACCGCGACACCGCGGTGG 1555  
QY 541 ATTTTCGCGCTGGCGGTTCGCGGCAAGCTGATCTGCTCCGCTGGGTGGGTTTCGCG 600  
Db 1556 ATTTTCGCGCTGGCGGTTCGCGGCAAGCTGATCTGCTCCGCTGGGTGGGTTTCGCG 1615  
QY 601 GCAGCAGCATTTGTCACCGCGCTGTCACCGCCCAAGTGTGGCGTGGATCAACGTCGTC 660  
Db 1616 GCAGCAGCATTTGTCACCGCGCTGTCACCGCCCAAGTGTGGCGTGGATCAACGTCGTC 1675  
QY 661 GTGGCAGTGTGATGACCGCATTCGCGCATCAAACTGATGTTGATGGGTAG 711  
Db 1676 GTGGCAGTGTGATGACCGCATTCGCGCATCAAACTGATGTTGATGGGTAG 1726

RESULT 6  
AP005277/c  
LOCUS  
DEFINITION Corynebacterium glutamicum ATCC 13032 DNA linear BCT 08-AUG-2002  
4/10  
ACCESSION AP005277 BA000036  
VERSION AP005277.1 GI:21323710  
KEYWORDS  
SOURCE  
ORGANISM Corynebacterium glutamicum ATCC 13032  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE  
1  
Nakagawa, S.  
Complete genomic sequence of Corynebacterium glutamicum ATCC 13032  
Unpublished  
2 (bases 1 to 333150)  
Nakagawa, S.  
Direct Submission  
TITLE Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.  
JOURNAL Ltd., Tokyo Research Laboratories; 3-6-6, Aza-hi-machi, Machida,  
Tokyo 194-8533, Japan (E-mail: snakagawa@xanagen.com,  
Tel: 81-44-829-3031, Fax: 81-44-813-1651)  
This sequence is conducted by collaboration of Kyowa Hakko Kogyo  
Co. Ltd. And Kitasato University.  
FEATURES  
Location/Qualifiers  
1..333150  
/organism="Corynebacterium glutamicum ATCC 13032"  
/mol\_type="genomic DNA"  
/strain="ATCC 13032"  
/db\_xref="taxon:196627"  
/note="ATCC 13032"  
complement (63..1292)  
/gene="Cg10944"  
complement (63..1292)  
/gene="Cg10944"  
/codon\_start=1  
/transl\_table=11  
/product="Hypothetical membrane protein"  
/protein\_id="BAB98337.1"  
/db\_xref="GI:21323711"  
/translation="MTNPTERNARLLIHWGLQNIQDQIVAAKTVLPWLLQAAGPG  
FLIALVPIREAGMLPQRAITGVWVLRQTSKRVWVIGNSQFVSALGIVAAFLRG  
WALGTITVLIALSLFRSMCSIASKDVQKVIKSKRGVLTVGRATVIGVWGLVAGL  
ATAIFLGSHPFTRVLAIVASSFSMLFASIVFARIEYAKPATPKNAPSANPWARCI  
AALKDKAFRRFVLVSMWLTALSTAFIVALAASGNSIDSLGFFLIASGLASWVG  
RISGWSHSHSKNMAGGALFGSIVLILVLSAFAPAQINTLVFPLSPFLITLAHTA  
IRVAKTIVMDMAEGDQTRIVVAQNTAMGVLLIVGALSGFIAPENAAALLFLAII  
GLLTGISARGLEKVSAG"

gene  
CDS

gene  
CDS  
1382..2917  
/gene="Cg10945"  
1382..2917  
/gene="Cg10945"  
/codon\_start=1  
/transl\_table=11  
/product="Putative multicopper oxidases"  
/protein\_id="BAB98338.1"  
/db\_xref="GI:21323712"  
/translation="MPEINRRTFKGAGVLAATVCAQVLVACSSDDVRVCGEPRTL  
PIPADLREGSSVHFALAGTGESQLPDDVTIKWGFNGHLGPTLVVKGGDDVHV  
DVNNLDMTIVWHMKLPALADGSPHSPGQTSPTMTVANDAATLWHPHTHG  
LTCFLAYEGALMIIVDEADTKLDRLEYGVDIPIVLMDRFLEDGSLDEEDLPDL  
BLQDTPFANGITNAHFADATRRVFRVLNGSNRMYNLATSDRTTQVTLASDGLLD  
EGDRTTLAIGPERWEIIVLEGEVDVTLSESVGREDNVGVPDDFEVDFRGMSSDFOL  
LITGPSDDAAQAPALPGVLAKSTPEPVIDATERTEIMNTFSINDLQMDQKRVDDVAIL  
HQQRVIVTNDSDMPHFHVDARFKVLFGFTDVELFNDGKWDVGLVPGATATID  
AVEFGYDPDPFMYTHCMLYHEDQGMQGVFVIEPGEPAALVIGSGSSIDSAGG  
HAH"  
3063..3983  
/gene="Cg10946"  
3063..3983  
/gene="Cg10946"  
/note="PF00005:ABC transporter  
TIGR00960:3a0501s02: Type II (General) Secretory Pathway  
(TISP) Family protein  
TIGR00968:3a0106s01: sulfate transport system permease  
protein  
TIGR01166:cbiO: cobalt transport protein ATP-binding  
subunit  
TIGR01184:ntrCD: nitrate transport ATP-binding subunits C  
and D  
TIGR01186:proV: glycine betaine/L-proline transport ATP  
binding subunit  
TIGR01187:pocA: spermidine/putrescine ABC transporter  
ATP-binding subunit  
TIGR01188:drdA: daunorubicin resistance ABC transporter  
ATP-binding subunit  
TIGR01189:ccmA: heme exporter protein CcmA"  
/codon\_start=1  
/transl\_table=11  
/product="ABC-type transporter, ATPase component"  
/protein\_id="BAB98339.1"  
/db\_xref="GI:21323713"  
/translation="MTTTPAIDVTLVRYGDTAVKGLNHFVORGVFLGLTGAG  
KSTLVEIGLSPSSSTVIRISGLDPVADRAILRPELQMLQSGELSQLTVAETMDM  
WHGTCTVPRATKQVLDVLRHENVKVGALSGEORLDLACALLGDSILFLDEPT  
TLGDESRHWQLLLDLKQGVMTMLTTHYLEAEFLCDRIAMNAGEIAVEGTLDE  
LVAREKSIISFVLRGQGVLPVLSGABIIRDNNHVRITATTLQOHTLEILTAAETGI  
ALEGFAKPATLESVFMDSLENTSLQTA"  
4003..4836  
/gene="Cg10947"  
4003..4836  
/gene="Cg10947"  
/codon\_start=1  
/transl\_table=11  
/product="ABC-type transporter, permease components"  
/protein\_id="BAB98340.1"  
/db\_xref="GI:21323714"  
/translation="MTTSTARGLQCHATPERKTSLEKAEWLQPRNKILLFM  
ATVPVGPLLLFLIGNGAASNSDFVMTLLVQVITVLSMATRRDRLVKR  
LRTGEARDIDIGALCPGALLITFTVPIPLMLVGPAPINLVIPVAILGLLL  
CSALAMTSGTFRNAEAQMTSPVFMALMGIGLSIRFVFGSDIVADILATPFAAIS  
DLVQIGWAGATFADSVGVGEANFAGIFQDMLPLIGLILAAATAAANRYRWDSY  
R"  
4879..6138  
/gene="Cg10948"  
4879..6138  
/gene="Cg10948"  
/EC\_number="2.7.3.-"  
/codon\_start=1  
/transl\_table=11  
/product="Two-component system, sensory transduction

histidine kinases"  
/protein\_id="BAB98341.1"  
/db\_xref="GI:21323715"  
/translation="NMKQFTAGTARFSGVSLASPVTPAPNSWRRLNPKAYTL  
VTRSLQAI PVLLSAYFLGVANAGTINPSFVLLGSFVILLIIVLVVEYQPSLMS  
HRRSVQPFPTGLVNLGVVSVLQIPLGLMSDNTRATLITLTCVFLSLIAYI  
PMWRVWMLMSAVLWMTSTTDYLSALWVPIPLMAGTVRLSVMTVDVMKVERS  
RELEASELVEERLRPAQELHDTLQGLHMSVKSSELALAKRGDORLELRELK  
LRTSMSEMDVSVGYRTNLEIEGAKSLADAHILHSVIGTTSQVSPARELCAM  
LYREATTNLRHSDATDATTLSSTEVMDNNGVNDIGRLSGLSALRSRBSAGMTL  
IVSREDDQFVRLINAPANTPAEKA"

6139..6750  
/gene="Cg10949"  
6139..6750  
/gene="Cg10949"  
/note="PF00072:Response regulator receiver domain  
PF00196:Bacterial regulatory proteins, luxR family"  
/codon\_start=1  
/transl\_table=11  
/product="two-component system, response regulators  
consisting of a CheY-like receiver domain and a HTH  
DNA-binding domain"  
/protein\_id="BAB98342.1"  
/db\_xref="GI:21323716"  
/translation="MISIADEALIASLALSLLEPDLVDVPTAGSGELIETWA  
DPNSTVCDVLDLQGGIDGIDATRLNETTPELAVLIVTSHAPROLKEALAAVLG  
FLPKSTADEFATIRVHAGRRYIDPELAAMTISAGESPLNREERVELLAGQSLA  
EELAVAAHAPGTTNRYLSQAMTKVQGNRFEAPTRARELGM"  
complement(6781..6987)  
/gene="Cg10950"  
complement(6781..6987)  
/gene="Cg10950"  
/codon\_start=1  
/transl\_table=11  
/product="Hypothetical protein"  
/protein\_id="BAB98343.1"  
/db\_xref="GI:21323717"  
/translation="MKHYSMVFGAGPLNTPHTKEVFMVDIVSVNHFWSWSTNNVWN  
MTGLPRETFVPAFVGSTVLAVLSF"

7048..8205  
/gene="Cg10951"  
7048..8205  
/gene="Cg10951"  
/codon\_start=1  
/transl\_table=11  
/product="tRNA-processing ribonuclease BN"  
/protein\_id="BAB98344.1"  
/db\_xref="GI:21323718"  
/translation="MTAMLCSPFSGVDYDMNGVQPOEHLDTLIAADPHGNPEN  
SGDKERLNPQMKVALNRTVDVFPDGLDLAALLTFSSILSIAPVLLGVSVITIF  
LASDTEILNVRDVNGYVPEQSHVNGVIDSIAGSAAAGVGVAVGITALTWTS  
AYVAFSRCANAVYGRSGRLIKRNAMLEFLNALLGIIILVSWLNELTMGIF  
APIAEPLHNLVSLFDREMPFIWVRFVIVGVLLFMVAILIYWAENRPFRL  
SLGSFLAIVGILLAGVGLNFYFTLPAFSSGAVGSLAVFTALVFNICLIIGLKID  
VEISRAKQLQAGPAEDYSLVPPRSIEKVKMKQKQORLMDQAAAREEN"

8672..10114  
/gene="Cg10952"  
8672..10114  
/gene="Cg10952"  
/note="PF00005:ABC transporter  
PF00664:ABC transporter transmembrane region."  
/codon\_start=1  
/transl\_table=11  
/product="ABC-type multidrug/protein/lipid transport  
system, ATPase component"  
/protein\_id="BAB98345.1"  
/db\_xref="GI:21323719"  
/translation="NMRTHTASSILRMRIRKQKVAFAFFLGMWOLSEALVPIAI  
GLIVDAVLTDRLRLVGLVAFVVLVSVLSYRFSRNLNRAVNFESHARLVEVAD  
HALKNDLPRMLVPEVMSKSTADSDSTRIFGQIGTGSAAATGELGAATYLLISDLV  
GLVLVLPVPIISGVWASAKGISKESVTOEKLAGQASQDMMGLRVKAIIGGERH  
AVKTEKASQASAAVDTAVASGVKAGISLAIANLAAVLLAGRVVTTGELGPQ  
LTAIVGAVVILSRPILLSNSINASAIAGHAERVAENFLNDESQAQYESSETINDGE

FLVIVVPASTPLPHGDNILAPHAADIEFTLRNLSMHEHNVBIDPOVIRASGLTDI  
IEVDGLDAPVRDTSNLSGQQRVALRALHADAEVLIVMDPISAVDSVTEVSTAQ  
IKQLRAGKATVIVVSSPAFYNLADRVISHV"

10107..11843  
/gene="Cg10953"  
10107..11843  
/gene="Cg10953"

Query Match 100.0%; Score 711; DB 1; Length 333150;  
Best Local Similarity 100.0%; Pred. No. 5.9e-169;  
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTGATCATGAAATCTTCAATACAGTCTGCTTTTGGGGCCAGTCTTTTACTCTCC 60  
Db 329003 ATGTGATCATGAAATCTTCAATACAGTCTGCTTTTGGGGCCAGTCTTTTACTCTCC 328944

Qy 61 ATCCGACCGCAGAAATGTAAGTCAATTAACAAAGAAATTAAGCGCGGAAGGACTCATTTGG 120  
Db 328943 ATCCGACCGCAGAAATGTAAGTCAATTAACAAAGAAATTAAGCGCGGAAGGACTCATTTGG 328884

Qy 121 GTTCTTCTCGTGTGTTTAAATTTCTGACGCTTTTGTTCATCGCGCGCACCTTGGGGGTT 180  
Db 328883 GTTCTTCTCGTGTGTTTAAATTTCTGACGCTTTTGTTCATCGCGCGCACCTTGGGGGTT 328824

Qy 181 GATCTTTTGTCCAAATGCGCGCGCGCATGCTCGATATTAAGCTGGGCTGGCATCGCT 240  
Db 328823 GATCTTTTGTCCAAATGCGCGCGCGCATGCTCGATATTAAGCTGGGCTGGCATCGCT 328764

Qy 241 TACCTGTATGTTTTCGCTCATCGCAGCGAAAGCGCATGACAAACAAAGGTGGAAGCG 300  
Db 328763 TACCTGTATGTTTTCGCTCATCGCAGCGAAAGCGCATGACAAACAAAGGTGGAAGCG 328704

Qy 301 CCACAGATCATTTGAGAAACAGAACCAACCGTCCCGATGACAGCGCTTTGGGGCGTTTCG 360  
Db 328703 CCACAGATCATTTGAGAAACAGAACCAACCGTCCCGATGACAGCGCTTTGGGGCGTTTCG 328644

Qy 361 GCGGTGGCCACTGACAGCGCGCAACCGGTCGGGTGGAGGTGAGCGTGCATTAAGCAGCGG 420  
Db 328643 GCGGTGGCCACTGACAGCGCGCAACCGGTCGGGTGGAGGTGAGCGTGCATTAAGCAGCGG 328584

Qy 421 GTTTCGGTAAAGCCCATGTTGATGCGCAATCGTCTGACCTGTTGAACCGCAATCGGTAT 480  
Db 328583 GTTTCGGTAAAGCCCATGTTGATGCGCAATCGTCTGACCTGTTGAACCGCAATCGGTAT 328524

Qy 481 TTGACAGCGCTTTGTGTTATTCGGCGCGGTTCGGCGCGCAATACGGCGACACCGCGGTGG 540  
Db 328523 TTGACAGCGCTTTGTGTTATTCGGCGCGGTTCGGCGCGCAATACGGCGACACCGCGGTGG 328464

Qy 541 ATTTTCGGCGTGGCGGTTTCGGCGCAAGCTGATCTGCTCCCGCTGGTGGGTTTCGGC 600  
Db 328463 ATTTTCGGCGTGGCGGTTTCGGCGCAAGCTGATCTGCTCCCGCTGGTGGGTTTCGGC 328404

Qy 601 GCACGACGATTTGTCACGCCCGCTGTCCAGCCCCCAAGGTGCGCGTGGATCAACGTCGTC 660  
Db 328403 GCACGACGATTTGTCACGCCCGCTGTCCAGCCCCCAAGGTGCGCGTGGATCAACGTCGTC 328344

Qy 661 GTGCGAGTTGTGATGACCGCATGCGCATCAACTGATGTTGATGGGTTAG 711  
Db 328343 GTGCGAGTTGTGATGACCGCATGCGCATCAACTGATGTTGATGGGTTAG 328293

RESULT 7  
AX127147/c  
LOCUS AX127147 349980 bp DNA linear PAT 11-MAY-2001  
DEFINITION Sequence 7063 from Patent EP1108790.  
ACCESSION AX127147 AX114121  
VERSION AX127147.1 GI:14041135  
KEYWORDS Corynebacterium glutamicum  
SOURCE Corynebacterium glutamicum  
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
REFERENCE 1

AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.  
TITLE Novel polynucleotides  
JOURNAL Patent: EP 1108790-A 7063 20-JUN-2001;  
KYOWA HAKKO KOGYO CO., LTD. (JP)  
FEATURES Location/Qualifiers  
source 1..349980  
/organism="Corynebacterium glutamicum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1718"  
/note="Seq 1 to long (3.309.400) split in 11, seq 7063  
1.200.001 1.549.980"  
ORIGIN  
Query Match 100.0%; Score 711; DB 6; Length 349980;  
Best Local Similarity 100.0%; Pred. No. 5.9e-169;  
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGTGATCATGGAATCTTCATTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTGCC 60  
Db 128953 ATGGTGATCATGGAATCTTCATTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTGCC 128894  
QY 61 ATCGACCGCAGATGTACTGTGATTAAACAAGGAATTAAGCGGAAGACTCATTTGCG 120  
Db 128893 ATCGACCGCAGATGTACTGTGATTAAACAAGGAATTAAGCGGAAGACTCATTTGCG 128834  
QY 121 GTTCTTCTGCTGTTTAAATTTCTGACCTCTTTTGTTCATCGCGGACCTTTGGCGTT 180  
Db 128833 GTTCTTCTGCTGTTTAAATTTCTGACCTCTTTTGTTCATCGCGGACCTTTGGCGTT 128774  
QY 181 GATCTTTTGTCAATGCGCGCGCATGCTGCTGATATTATGCGCTGGGCTGGCATCGCT 240  
Db 128773 GATCTTTTGTCAATGCGCGCGCATGCTGCTGATATTATGCGCTGGGCTGGCATCGCT 128714  
QY 241 TACCTGTTATGTTTTCGCTGATGCGAGAAAGAGCCCATGACAAACAAGGTGGAAGCG 300  
Db 128713 TACCTGTTATGTTTTCGCTGATGCGAGAAAGAGCCCATGACAAACAAGGTGGAAGCG 128654  
QY 301 CCACAGATCATTTGAAGAAACAGAACCGTGGCCGATGACACGCTTTTGGCGGTTTCG 360  
Db 128653 CCACAGATCATTTGAAGAAACAGAACCGTGGCCGATGACACGCTTTTGGCGGTTTCG 128594  
QY 361 GCGGTGGCCACTGACACGCGCGCGGTGGGCTGAGGTGAGTGGATTAAGCAGCGG 420  
Db 128593 GCGGTGGCCACTGACACGCGCGCGGTGGGCTGAGGTGAGTGGATTAAGCAGCGG 128534  
QY 421 GTTTGGGTAAGCCCATTTGATGGAATCTGCTGATCTGCTGATGATGATGATGATGAT 480  
Db 128533 GTTTGGGTAAGCCCATTTGATGGAATCTGCTGATCTGCTGATGATGATGATGATGAT 128474  
QY 481 TTGGACGCGTTTGTGTTATCGCGCGGTGGGCTGAGTGGATGAGTGGATGAGTGGAT 540  
Db 128473 TTGGACGCGTTTGTGTTATCGCGCGGTGGGCTGAGTGGATGAGTGGATGAGTGGAT 128414  
QY 541 ATTTTCGCGCTGGCGGTTTCGCGGCAAGCCTGATCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 128413 ATTTTCGCGCTGGCGGTTTCGCGGCAAGCCTGATCTGCTGCTGCTGCTGCTGCTGCTG 128354  
QY 601 GCGACGAGTGTGACGCGCGCTGTCAGCGCCCAAGGTGAGGTGAGGTGAGGTGAGGTGAG 660  
Db 128353 GCGACGAGTGTGACGCGCGCTGTCAGCGCCCAAGGTGAGGTGAGGTGAGGTGAGGTGAG 128294  
QY 661 GTGCGAGTGTGATGACCGCATTTGGCCATCAAACTGATGATGATGATGATGATGATGAT 711  
Db 128293 GTGCGAGTGTGATGACCGCATTTGGCCATCAAACTGATGATGATGATGATGATGATGAT 128243  
RESULT 8  
AX123539  
LOCUS AX123539 708 bp DNA linear PAT 11-MAY-2001  
DEFINITION Sequence 3455 from Patent EP1108790.  
ACCESSION AX123539  
VERSION AX123539.1 GI:14041027

KEYWORDS Corynebacterium glutamicum  
SOURCE Corynebacterium glutamicum  
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriaceae; Corynebacteriidae; Corynebacterium.  
REFERENCE 1  
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.  
TITLE Novel polynucleotides  
JOURNAL Patent: EP 1108790-A 3455 20-JUN-2001;  
KYOWA HAKKO KOGYO CO., LTD. (JP)  
FEATURES Location/Qualifiers  
source 1..708  
/organism="Corynebacterium glutamicum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1718"  
ORIGIN  
Query Match 99.6%; Score 708; DB 6; Length 708;  
Best Local Similarity 100.0%; Pred. No. 1.5e-168;  
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGCTGATCATGGAATCTTCATTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTGCC 60  
Db 1 ATGCTGATCATGGAATCTTCATTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTGCC 60  
QY 61 ATCGACCGCAGATGTACTGTGATTAAACAAGGAATTAAGCGGAAGACTCATTTGCG 120  
Db 61 ATCGACCGCAGATGTACTGTGATTAAACAAGGAATTAAGCGGAAGACTCATTTGCG 120  
QY 121 GTTCTTCTGCTGTTTAAATTTCTGACCTCTTTTGTTCATCGCGGACCTTTGGCGTT 180  
Db 121 GTTCTTCTGCTGTTTAAATTTCTGACCTCTTTTGTTCATCGCGGACCTTTGGCGTT 180  
QY 181 GATCTTTTGTCAATGCGCGCGCATGCTGCTGATATTATGCGCTGGGCTGGCATCGCT 240  
Db 181 GATCTTTTGTCAATGCGCGCGCATGCTGCTGATATTATGCGCTGGGCTGGCATCGCT 240  
QY 241 TACCTGTTATGTTTTCGCTGATGCGAGAAAGAGCCCATGACAAACAAGGTGGAAGCG 300  
Db 241 TACCTGTTATGTTTTCGCTGATGCGAGAAAGAGCCCATGACAAACAAGGTGGAAGCG 300  
QY 301 CCACAGATCATTTGAAGAAACAGAACCGTGGCCGATGACACGCTTTTGGCGGTTTCG 360  
Db 301 CCACAGATCATTTGAAGAAACAGAACCGTGGCCGATGACACGCTTTTGGCGGTTTCG 360  
QY 361 GCGGTGGCCACTGACACGCGCGCGGTGGGCTGAGGTGAGTGGATTAAGCAGCGG 420  
Db 361 GCGGTGGCCACTGACACGCGCGCGGTGGGCTGAGGTGAGTGGATTAAGCAGCGG 420  
QY 421 GTTTGGGTAAGCCCATTTGATGGAATCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTG 480  
Db 421 GTTTGGGTAAGCCCATTTGATGGAATCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTG 480  
QY 481 TTGGACGCGTTTGTGTTATCGCGCGGTGGGCTGAGTGGATGAGTGGATGAGTGGAT 540  
Db 481 TTGGACGCGTTTGTGTTATCGCGCGGTGGGCTGAGTGGATGAGTGGATGAGTGGAT 540  
QY 541 ATTTTCGCGCTGGCGGTTTCGCGGCAAGCCTGATCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 541 ATTTTCGCGCTGGCGGTTTCGCGGCAAGCCTGATCTGCTGCTGCTGCTGCTGCTGCTG 600  
QY 601 GCGACGAGTGTGACGCGCGCTGTCAGCGCCCAAGGTGAGGTGAGGTGAGGTGAGGTGAG 660  
Db 601 GCGACGAGTGTGACGCGCGCTGTCAGCGCCCAAGGTGAGGTGAGGTGAGGTGAGGTGAG 660  
QY 661 GTGCGAGTGTGATGACCGCATTTGGCCATCAAACTGATGATGATGATGATGATGATGAT 708  
Db 661 GTGCGAGTGTGATGACCGCATTTGGCCATCAAACTGATGATGATGATGATGATGATGAT 708  
RESULT 9  
BD165656

D	B	361	GCGGTGGCCACTGACACGCCGAACCGGTGGGGTGGAGGTGCAGCTGCATTAAGCAGCGG	420
Q	Y	421	GTTTGGGTAAAGCCCATTGTTGATGGCAATCGTGCTGACCTGGTTGAACCCGAATGCGTAT	480
D	B	421	GTTTGGGTAAAGCCCATTGTTGATGGCAATCGTGCTGACCTGGTTGAACCCGAATGCGTAT	480
Q	Y	481	TTCGACGGCTTTGTGTTATCGGCGCGCTCGGCGCGCATACGGCGCACCGGACGGTGG	540
D	B	481	TTCGACGGCTTTGTGTTATCGGCGCGCTCGGCGCGCATACGGCGCACCGGACGGTGG	540
Q	Y	541	ATTTTCGCGCTGGCGCGCTTCGGCGCAAGCTCGATCTGGTTCCTCGCTGGTGGGTTTCGGC	600
D	B	541	ATTTTCGCGCTGGCGCGCTTCGGCGCAAGCTCGATCTGGTTCCTCGCTGGTGGGTTTCGGC	600
Q	Y	601	GCAGCAGCATTTGTCACGCCCGCTGTCACGCCCCAAAGCTGTGGTTCCTCGCTGGTGGGTTTCGGC	660
D	B	601	GCAGCAGCATTTGTCACGCCCGCTGTCACGCCCCAAAGCTGTGGTTCCTCGCTGGTGGGTTTCGGC	660
Q	Y	661	GTGGCAGTTGTGATGACCGCATTTGGCCATCAAATCGATGTTGATGGT	708
D	B	661	GTGGCAGTTGTGATGACCGCATTTGGCCATCAAATCGATGTTGATGGT	708
<hr/>				
RESULT 10				
AX643030				
LOCUS				
DEFINITION Sequence 9 from Patent EP1266966.				
ACCESSION AX643030				
VERSION AX643030.1 GI:28550160				
KEYWORDS Corynebacterium glutamicum				
SOURCE Corynebacterium glutamicum				
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
Corynebacterineae; Corynebacteriaceae; Corynebacterium.				
REFERENCE 1 Gunji, Y. and Yasueda, H.				
AUTHORS Method for producing L-lysine or L-arginine by using methanol				
TITLE assimilating bacterium				
JOURNAL Patent: EP 1266966-A 9 18-DEC-2002;				
Ajinomoto Co., Inc. (JP)				
FEATURES Location/Qualifiers				
source 1..712 linear PAT 24-FEB-2003				
/organism="Corynebacterium glutamicum"				
/mol_type="unassigned DNA"				
/db_xref="taxon:1718"				
CDS 1..375				
/note="unnamed protein product"				
/codon_start=1				
/translation="WTMEIPITGLLGASLLLSIGPQNVLVIKQIKREGHIAVLE				
CLISDVPFLFAGTLGVDLNNAAPIVLDIMRWGGIAYLLWFAVMAAKDMTKNEAP				
IIIEETPTWDDTPPLGVFGGH"				
<hr/>				
ORIGIN				
Query Match 98.0%; Score 696.8; DB 6; Length 712;				
Best Local Similarity 99.6%; Pred. No. 1.1e-165;				
Matches 709; Conservative 0; Mismatches 2; Indels 1; Gaps 1.1;				
<hr/>				
Q	Y	1	ATGTGATCATGGAATCTTCATACAGTCTGCTTTTGGGGGCGAGTCTTTTACTGTCC	60
D	B	1	ATGTGATCATGGAATCTTCATACAGTCTGCTTTTGGGGGCGAGTCTTTTACTGTCC	60
Q	Y	61	ATCGGACCCGACATGTACTGTGTGATTAACAGGAATAAGCGCAAGGACTCATTTGGC	120
D	B	61	ATCGGACCCGACATGTACTGTGTGATTAACAGGAATAAGCGCAAGGACTCATTTGGC	120
Q	Y	121	GTTCCTTCGTGTGTTAAATTCACAGTCTTTTTTGTTCATCGCGCGACCTTGGCGTT	180
E	S	131	ATTCCTTCGTGTGTTAAATTCACAGTCTTTTTTGTTCATCGCGCGACCTTGGCGTT	180

181 GATCTTTGTCATGCGCGCGATCGTCTGATATTATGCTGCGGTGGGTCGCTGCT 240  
 181 GATCTTTGTCATGCGCGCGCGATCGTCTGATATTATGCTGCGGTGGGTCGCTGCT 240  
 241 TACCTGTTATGCTTTGCTGCTGATGCGCGCGCGATGCGCGCGCGCGCGCGCG 300  
 241 TACCTGTTATGCTTTGCTGCTGATGCGCGCGCGCGATGCGCGCGCGCGCGCG 300  
 301 CCACGATCATTGAAAGAAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 359  
 301 CCACGATCATTGAAAGAAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
 360 GCGGTGGCG 419  
 361 GCGGTGGCG 420  
 420 GGTGTTGGTAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 479  
 421 GGTGTTGGTAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
 480 TTTGGACCG 539  
 481 TTTGGACCG 540  
 540 GATTTTCG 599  
 541 GATTTTCG 600  
 600 CGGACGACATGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 659  
 601 CGGACGACATGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660  
 660 CTTGCG 711  
 661 CTTGCG 712

RESULT 11  
 E54483  
 LOCUS 1568 bp DNA linear PAT 27-AUG-2002  
 DEFINITION Heat-resistant lysine biosynthesis enzyme gene of thermophilic  
 coryneform bacterium.  
 ACCESSION E54483  
 VERSION 1  
 KEYWORDS JP 2001120270-A/7.  
 SOURCE Corynebacterium thermoaminogenes  
 ORGANISM Corynebacterium thermoaminogenes  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacteriaceae; Corynebacteriidae; Corynebacterium.  
 REFERENCE 1 (bases 1 to 1568)  
 AUTHORS Itaya, M., Kimura, E., Kawara, Y. and Sugimoto, S.  
 TITLE Heat-resistant lysine biosynthesis enzyme gene of thermophilic  
 coryneform bacterium  
 JOURNAL Patent: JP 2001120270-A 7 08-MAY-2001;  
 COMMENT AUINOMOTO CO INC  
 OS Corynebacterium thermoaminogenes  
 PN JP 2001120270-A/7  
 PD 08-MAY-2001  
 PF 01-NOV-1999 JP 1999311148  
 PI MINOKU ITAYA, EICHIRO KIMURA, YOSHITO KAWARA, SHINTACHI SUGIMOTO PC  
 C12N15/09/(C12N15/09, C12N15/00, C12N15/00, C12N15/00, C12N15/00) CC  
 FH Key Location/Qualifiers  
 FT CDS (778)..(1470).  
 FEATURES  
 source 1..1568  
 /organism="Corynebacterium thermoaminogenes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:161484"  
 ORIGIN  
 Query Match 44.7%; Score 317.6; DB 6; Length 1568;  
 Best Local Similarity 67.7%; Pred. No. 1.5e-69;  
 Matches 481; Conservative 0; Mismatches 214; Indels 15; Gaps 2;

1 ATGCTGATCATGGAATCTTATACAGTCTGCTTTGGGGGCCAGTCTTTTACTGCTC 60  
 778 ATGCGGACATGGAATCTTATACAGTCTGCTTTGGGGGCCAGTCTTTTACTGCTC 837  
 61 ATCGGACCGCGAATGATCTGCTGATTAACAGAGAAATTAAGCGGAGGACTCATTTGG 120  
 838 ATCGGCGCACAGATGCTCTGCTGATCAACAGGGCATCAACAGCGGAGGCGATCAGCGC 897  
 121 GTTCTTCTGCTGCTTAAATTTCTGAGCTCTTTTCTGATCGCGGACCTTTGGGCGTT 180  
 898 GTCATCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957  
 181 GATCTTTGCTCAATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 958 GGCCTGATCTCGACACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017  
 241 TACCTGTTATGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 1018 TACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
 301 CCACGATCATTGAAAGAAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
 1078 ACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122  
 361 GCGGTGGCG 420  
 1123 GCGGCG 1182  
 421 GTTGGGTAAGCG 480  
 1183 GTCTGGGTCAGCG 1242  
 481 TTGCG 540  
 1243 CTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1302  
 541 ATTTTCG 600  
 1303 ATCTTCGCTGCG 1362  
 601 GCGGACGATTTGTCAGCG 660  
 1363 GCGGCG 1422  
 661 GTGCGAGTTGTCAGCG 710  
 1423 GTGCG 1472

AB083133 1771 bp DNA linear BCT 06-APR-2002  
 DEFINITION Corynebacterium efficiens lysG, lysE genes for Lysine export  
 transcriptional regulatory protein, Lysine exporter protein,  
 complete cds.  
 ACCESSION AB083133  
 VERSION 1  
 KEYWORDS Corynebacterium efficiens  
 SOURCE Corynebacterium efficiens  
 ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacteriaceae; Corynebacteriidae; Corynebacterium.  
 REFERENCE 1  
 AUTHORS Itaya, H., Kimura, E., Kawahara, Y. and Sugimoto, S.  
 TITLE lysG, lysE of Corynebacterium efficiens  
 JOURNAL Published Only in Database (2002)  
 REFERENCE 2 (bases 1 to 1771)  
 AUTHORS Itaya, H., Kimura, E., Kawahara, Y. and Sugimoto, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-APR-2002) Hiroshi Itaya, AJINOMOTO CO., INC.  
 fermentation & Biotechnology Laboratories, 1-1, Suzuki-cho,  
 Kawasaki-Ku, Kawasaki, Kanagawa 210-8661, Japan



(E-mail:hiroshi\_itaya@ajinomoto.com, Tel:81-44-244-7123(ex.4146),  
Fax:81-44-222-0129)

```

FEATURES
    Location/Qualifiers
        1..1771
            /organism="Corynebacterium efficiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:152794"
            /complement(47..919)
            /gene="lysG"
            complement(47..919)
            /gene="lysG"
            /codon_start=1
            /transl_table=11
            /product="lysine export transcriptional regulatory
            protein"
            protein
            /protein_id="BAB88826.1"
            /db_xref="GI:20065739"
            /translation="MNPILHDLTIIIDEGSPENASIALSISPSAVSORIKALEKSVG
            RVLSRTOPAVATEAGEVLVQARPMALLOAETREGLAERLDEIPLTVAINADSLSTM
            FPPVFAVHGAVTILRLVEDEANTLSLRGSLVLCATVREADPVGCEVLRIGVMR
            HLPVATPELRARYTVDDQDWMFVLRFGNDVLDGRVDFDGVAVARRVSVVPS
            AEGFGEAVRLGLGMLPEAQAAPLAAGDVQDLEKVVDTPLVQWRLESRLRLARL
            TDVVDAARAGLRT"
            981..1676
            /gene="lysE"
            981..1676
            /gene="lysE"
            /codon_start=1
            /transl_table=11
            /product="lysine exporter protein"
            protein
            /protein_id="BAB88827.1"
            /db_xref="GI:20065740"
            /translation="MRDMEIFVTGLLGASLLAILAIGPQNVLVKQIKRGITAVILV
            CLISDVLTIGTIGVLGSDTAPILDLRWGCIAYLWFAVWMAARDALRAKTEVIF
            VHESEVAAASGSGVTKQRPRLITSTGRQVYRPMALVLTWLNFNATYDAFV
            FIGGVAQVGTGRIWIPAGATAAASLVWFLVGVGAALSRPLSPRVWRWINGVAV
            VLTLGLAVKLLWG"

ORIGIN
    Query Match 44.7%; Score 317.6; DB 1; Length 1771;
    Best Local Similarity 67.7%; Pred. No. 1.5e-69;
    Matches 481; Conservative 0; Mismatches 214; Indels 15; Gaps 2;

QY 1 ATGGTGATCATGAATCTTATTACAGGTCTGCTTTGGGGCCAGTCTTTTATCTGCC 60
DB 1111111111111111111111111111111111111111111111111111111
QY 981 ATGCGGACATGAAATCTTGTACCGGTTGTTGTTGGAGCCAGTCTGCTGTGCC 1040
DB 1111111111111111111111111111111111111111111111111111111
QY 61 ATCGGACCGCAGAAATGACTGTGATTAACAAGGAATTAAGCGCGAAGGACTATTGG 120
DB 1041 ATCGGCCACAGAAATCTCTGTGATCARACAGGGCATCAACGCGAGGCGATCAGGCC 1100
QY 121 GTTCTTCGTGTGTTAAATTTCTGACGTCTTTTGTTCATCGCCGGCACCTTGGGGTT 180
DB 1101 GTCATCATCGTCTGTCTGTCTCGAGCTGGTCTGTTTCACTCGGCACTCGGGGTC 1160
QY 181 GATCTTTTGGCCATCGCGCCGATCGTCTCGATATTATGCTGGGGTGGCATCGCT 240
DB 1161 GGCCTGATCTCGACACCCGCCGATCATTTCTGCACATCTGCTGGTGGCGCATCGCC 1220
QY 241 TACCTGTTATGTTTCCGTCATGCGCAGCGAAGACGCCCATGACAAACAAGGTGGAAGCG 300
DB 1221 TACCTGCTGTGGTTGCGGTGATGCGCGCGCGAGCGCTCGCGCCCGCACCGAGTA 1280
QY 301 CCACAGATCATTTAGAGAAACAGAACCAACACGTCGCCGATGACAGCCCTTGGCGGTTCG 360
DB 1281 ACCTTTGTC-----GAGCATTCGGAACCCGTTG---CGGACAGCTCCGCCCTCC 1325
QY 361 GCGGTGGCCACTGACACGCGCAACCGGCTGGGGTGGAGGTGAGCGTCCGATAGCAGCGG 420
DB 1326 GCGGGGGCGGTGACGAGAAACAGACCCCGGCTCCGATCATCATGACGACCCGGCAG 1385
QY 421 GTTTGGGTAAAGCCCATGTTGATGGCAATCGTCTGCACTGGTTGAACCCGATCGGAT 480
DB 1111111111111111111111111111111111111111111111111111111

```

```

Db 1386 GTCTGGTCCAGGCCCATCTCATGGCCATTGTGCTGACCTGGCTCAATCCCAATGCTAC 1445
QY 481 TTGAGCCGCTTTGTTTATTCGGCGGCGTTCGGCGGCACATACCGCGACACCGGCGTGG 540
DB 1446 CTGGATCCCTTCTCTTCATCGGTGGTGTGCGGAGCCAGTACGGGGAGACCGGTCGGTGG 1505
QY 541 ATTTTTCGCGCTGGCGCTTCGGCGCAAGCTGATCTGCTTCCCGCTGGTGGTGGTTCGGC 600
DB 1506 ATCTTCCTGCGGTGCTTTCGGCGCAGCTGTGCTTGTTCCTCCACTGGTGGTTCAGGC 1565
QY 601 GCACGACATTTGTACGCCCGCTGTCCAGCCCAAGAGTGTGGCGCTGGATCAAGTGTGC 660
DB 1566 GGGGCGGCACTGTGCGCTCCCTGTCCAGCGCGGCTGTGGCGCTGGATCAACATAGT 1625
QY 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAACTGATGTTGATGGGTTA 710
DB 1626 GTGGCCCTGTGCTCACCGGATTTGGCGTGAAGCTGATCTGTATGGGTTA 1675

RESULT 13
AP005218/c 308650 bp DNA linear BCT 10-JUL-2003
LOCUS Corynebacterium efficiens YS-314 DNA, complete genome, section
DEFINITION AP005218.1 GI:23493016
ACCESSION AP005218 BA000035
VERSION AP005218.1
KEYWORDS Corynebacterium efficiens YS-314
SOURCE Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
ORGANISM Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E.,
Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and
Gojobori, T.
TITLE Comparative Complete Genome Sequence Analysis of the Amino Acid
Replacements Responsible for the Thermotability of Corynebacterium
efficiens
JOURNAL Genome Res. 13 (7), 1572-1579 (2003)
MEDLINE 22723752
PUBMED 12840036
REFERENCE 2 (bases 1 to 308650)
AUTHORS Kawarabayasi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and
Director-General of Biotechnology Center.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-2002) Director-General of Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066 Japan
[E-mail:biomite.go.jp, Tel:81-3-3481-1933, Fax:81-3-3481-8424]
Kawarabayasi, Y. is officially affiliated with the National
Institute of Advanced Industrial Science and Technology, Tsukuba,
Ibaraki, 305-8566 Japan
Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the
National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan
Itoh, T. is at the Japan Biological Information Research Center,
Koto-ku, Tokyo, 135-0064 Japan
Yamagishi, A. is at Tokyo University of Pharmacy and Life Science,
Hachioji, Tokyo, 192-0392 Japan
Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co.,
Inc., Kawasaki, Kanagawa, 210-8681 Japan
The other authors are at the National Institute of Technology and
Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.
LOCATION/Qualifiers
    1..308650
    /organism="Corynebacterium efficiens YS-314"
    /mol_type="genomic DNA"
    /strain="YS-314"
    /db_xref="taxon:196164"
    complement(83..241)
    /note="CE1178"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"

CDS

```

/protein\_id="BAC17988.1"  
/db\_xref="GI:23493017"  
/translation="MTTGLILGTGPDAGLAKLGSIDIPRGVGIHRFRKPPVALQVR  
TGAGARD"  
365..742  
/note="CE1179"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17989.1"  
/db\_xref="GI:23493018"  
/translation="MEKQQSQKXQDFASVDHLSDDAAAMFIDNELSRGAMHRAHLH  
LVHCAECRQEIARQDVTDFVTECRNDNTAPSLKAKLMSLANECAPGCAADAAL  
QRPESFVAKVESVVRVRRTOGR"  
774..1271  
/note="CE1180, similar to AJ292256-1|CAC21724.1| percent  
identity: 26 in 155 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17990.1"  
/db\_xref="GI:23493019"  
/translation="MFSNVGGEILLFVIGLIVIGPERILPRLMEDLKAALVAARTAI  
NNAQNLDTFGEEDFDIRKPMQIAQIRMSGFKTAITRALFDGDSFLDSFPDKIM  
ADDTAGQARHQVAKTGTTPHEAEVVERPADRIDTQSPNTPKGNNDGKPPAGGYT  
WTDLI"  
complement (1353..2480)  
/note="CE1181, similar to AL390975-15|CAC01353.1| percent  
identity: 63 in 370 aa"  
/codon\_start=1  
/transl\_table=11  
/product="putative ATP-binding protein Mrp"  
/protein\_id="BAC17991.1"  
/db\_xref="GI:23493020"  
/translation="MMSQVTSASVSALSRVEDPEIGKPITELMQVKSVAIDGSDVQV  
EYITIAACPMKTIIVSNTAAIGEGVDQVTTDVMSDEQRRLVSLRGNTADP  
VYPFQAGSTRTVYVASKGVGKSVNTALATAAKGLSVGLDADIYGHVPGM  
LGSARPHQVDIMPPQAGKVMKISIAHFVEGNAPVVRGPMHRAIQOFLSDVFWG  
LDLILLDPPGDIATIAQLPNAELLIVTPQAAADVAERAGSISLQTNOKVA  
GVENMAMVPGSTMEVFGSGGKIAQRLSMLTGEDVPVLSGVPDPSLRVGGDV  
GNPTAISEPSPPTTAANAIDTIAKKKVSIVGKTLGLGVK"  
complement (2513..3271)  
/note="CE1182"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17992.1"  
/db\_xref="GI:23493021"  
/translation="MTTATPWADAFDHVGWDLFOILEVDHTERAVAHLOEYFATDPRT  
GPMYTGSLNRLGGGDAQPVRENIITRDDLLSLILDVVPVGLVQVQLLEPPEDAQSI  
DLEWVTRDYAVELPYDLEQILGILASIPTDVLGASKAAEHLDAENLLMTALRRNG  
PGPRVSKLIARPKLPFIQDSALKQNTNSGVFENFTVLRADKALSHLKG  
RTEAADSGNSALADLSIRVDFVVMADHKKI"  
complement (3282..3836)  
/note="CE1183, similar to AE007003-3|AAK45527.1| percent  
identity: 57 in 159 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17993.1"  
/db\_xref="GI:23493022"  
/translation="MADYSRDLDTPLVGTKKRRFKFDDTVGAVAEKVARFFGTGQY  
LFWQTFIWMINLGGFAMWDPYFILLNIAFSTQAAVAAPLILAAQNRQEDRK  
STILADRAEETKADTELTETELAGVLRVAVGEMVTRYLRHELDLRLGLELEAKL  
DAVARNDDPHILNAPTOCKVD"  
complement (3829..5124)  
/note="CE1184, similar to AE007003-4|AAK45528.1| percent  
identity: 52 in 417 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17994.1"  
/db\_xref="GI:23493023"

/translation="MRGITRLVAGRLAGMVRGPDVTIGRVDVVVVVIRSGQDSRV  
LGLVVLNNKRIPLPMLRTAAIEPEKVNLSGSVLSRSEFKARTGELOVIGDIITGK  
VHTDDPELQKHRAVETADVELEQTTTRDMWISVVALIGSRGFRPTLYIAPWSH  
VHGITTAGAGETQTAELIAGFEDMARADVAKELYGLPKFORDVAVAGELDDRADIL  
QELSDRQALIEITDIERAADVLEEMDPDAADLLGELPEDKAGVLELMDPEESAP  
VRLMGFNPDTVGMALMTEPLTMDPTTVAALAMARNPDLPTSLASLVFVVRPPTAT  
PTGKYLGCVLKQLLREPPSLVGILDPDLPPLVANSQETAARFFATFATVNLVCGFVL  
DENGHLIGAVAVDDLLDEMLPEDWRDAGIRPGVKETILG"  
5264..5782  
/note="CE1185, similar to AE007003-6|AAK45530.1| percent  
identity: 60 in 155 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17995.1"  
/db\_xref="GI:23493024"  
/translation="MSMQATSGGNNRAELGLKKRPEGVPVGSQTYEEARQAVDMLS  
DNEFFVABLTIIVGVDLMEVETVTRLTWGRVLAGGAAGMLGLFGLVMGLVSLWF  
SSLLTGLVMGVVFGVVLAAVFWASRGKRDFTSATQIVAGRYDILCRDPRAREARDMI  
ARNMGLGHPQ"  
5805..6500  
/note="CE1186"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17996.1"  
/db\_xref="GI:23493025"  
/translation="MTTGVKKALIGLVAGSLIVAGSGVVDNREFFEPHQVAREOK  
ENLYTAVDDPYLGHNVISLEYKVLNGRFAVIRHENDLIADFPDLNMGTSIDLIVG  
CTGALLEVNPVLAGELEEEYLAAVEAGETDINSGEWRDFTDALIGSLRLDITADP  
SNAQCGANGLPQNTIVPIYRNPTFSRADREILNWSAGAITTSELQELBEAKELGRS  
RELVKFVASKGG"  
6660..7373  
/note="CE1187, similar to AE007262-3|AAK65408.1| percent  
identity: 32 in 235 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17997.1"  
/db\_xref="GI:23493026"  
/translation="METILLGVLFVACVQGAIGFGLGMLAAPILALMRPOLLGLLI  
LLAFGSLATWARDGVEVPMVGVSSLLGRIPGSLVIGTWAVMPVAGUKIVLATAV  
ILGLTISLVGNSPGHGRNSFLAGSGGFLGTTAIGGPPALIMRSMSRERVGTLS  
VCFVIGSAISALTGALGMHMLQALVYAPAVIAGYFLSGVGVNKLNRRLIFLGS  
VTISLISGSMVITQAAGIF"  
7592..8149  
/note="CE1188"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17998.1"  
/db\_xref="GI:23493027"  
/translation="MSRGDFPAPFPADPCRGKTIKMQKMTKGAIATAAALLLV  
GGGTGFMWDEVTQSDQAIAGNLALSAADGKWTLNGDDVDVNVAVVVGDTLV  
FQGTISYTSQDNLSATLITLTPGSLGTATADAALATLNDTTITDILGELPSTIG  
SNTYTFRQVQSGVSGVSAISIKQA"  
complement (8190..9443)  
/note="CE1189, similar to AF024666-23|AAG03374.1| percent  
identity: 84 in 408 aa"  
/codon\_start=1  
/transl\_table=11  
/product="transposase"  
/protein\_id="BAC17999.1"  
/db\_xref="GI:23493028"  
/translation="MAAGPYSIDPTVYLLDELAQSPDLMRELMQGFINOILSTQADQ  
VGAAYATVSDRVNTNGRHRDPDVTGVTIDVAPKILRTGSPFDDMLLRTRAEER  
AUTVTIATYKLGKSTERNMDLVATLGNLSKQSVSMADQDQMWEEFRPRLDTG  
PLYVSCDALTMKVRGEGVVKTSVLLATGNVAGYRELLGKQVATASVPSVGFGR  
DLKARGLENYLVISDAHLGIOHAIGVEIPNMQRCRTHFSKLYGMWSKQWPTLS  
AMFHTIQQDPAQSVMGQVAVFCQKFPDVADYLEALDLAFATMYSKVMYKV  
WSNPTERLNEIRRTDVVGIFFPNRQAVVRLVAGVLAQBQDDMIQOKRYSLSLEQ  
TRAMTANIISDDITSEILIKDA"  
complement (11439..15143)

gene

```

CDS
/ gene="odha"
complement (11439..15143)

Query Match 44.7%; Score 317.6; DB 1; Length 308650;
Best Local Similarity 67.7%; Pred. No. 2.9e-69;
Matches 481; Conservative 0; Mismatches 214; Indels 15; Gaps 2;

QY 1 ATGGTGATCATGGAATCTTCATACAGGTCTGTTTGGGGCCAGCTTTTACTGTCC 60
Db 191614 ATGGGGACATGGAATCTTGTACCGGTTTGTGGGAGCAGCTCTCTGTGGCC 191555

QY 61 ATCGGACCGCAGATGTAAGTGAATTAAGCAAGAAATTAAGCGCAAGAGCTCATTTGG 120
Db 191554 ATCGGCCACAGATGCTCTGTGATCAACAGGCGATCAACGCGAGGCGATCAAGGCC 191495

QY 121 GTTCTCTGCTGTGTTAAATTTCTGACGCTCTTTTGTATCGCGGACCTTGGGGTT 180
Db 191494 GTCATCATCTGCTGTCTCTGCTGCGAGCTGCTGTGTTTCACTCGGACCTCGGGGTC 191435

QY 181 GATCTTTTGTCAATGCGCGCGATGCTGCTGATATATGCTGGGCTGGGTCATCGT 240
Db 191434 GGCCTGATCTTCGACACCGCCCGATCATCTTCGATCTGCTGCTGGGTCGATCGC 191375

QY 241 TACCTGTTATGTTTTCGCTCATGCGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 300
Db 191374 TACCTGCTGTGTTTTCGCTCATGCGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 191315

QY 301 CCACAGATCATGGAAGAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 360
Db 191314 ACCTTTGTC-----GAGCATTCGGAACCCGCTG---CGGACGCTGCGCTCC 191270

QY 361 GCGGTGGCCACTGACACGCGCAACCGGCTGCGGCTGAGGTGAGCTGATCAAGCAGCG 420
Db 191269 GCGGGGGGCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 191210

QY 421 GTTGGTAAAGCCCATGTTGATGCAATCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 191209 GTCTGGTCAAGCCCATGCTCATGGCCATGCTGCTGCTGCTGCTGCTGCTGCTG 191150

QY 481 TTGACCGGTTTGTGTTTATCGCGGCTGCGGCAATCAAGCAAGCAAGCAAGCAAG 540
Db 191149 CTGATCGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 191090

QY 541 ATTTTCCCGCTGCGGCTGCTGCGCAAGCTGATCTGTTTCCGCTGCTGCTGCTGCTG 600
Db 191089 ATCTTCTGCTGCGGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 191030

QY 601 GCAGGAGATGTCAGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTG 660
Db 191029 GCGGCGGCACTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 190970

QY 661 GTGGCAGTTGATGACGCAATGCGCATCAAGCAAGCAAGCAAGCAAGCAAGCAAG 710
Db 190969 GTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 190920

RESULT 14
BX248357/c 349535 bp DNA linear BCT 06-NOV-2003
LOCUS Corynebacterium diphtheriae gravis NCTC13129, complete genome;
DEFINITION segment 4/8.
ACCESSION BX248357 BX248353
VERSION BX248357.1 GI:38199912
KEYWORDS Complete genome.
SOURCE Corynebacterium diphtheriae
ORGANISM Corynebacterium diphtheriae
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1 (bases 1 to 349535)
AUTHORS Cerdano-Tarraga, A.M., Efstathiou, A., Dover, L.G., Holden, M.T.G.,
Pallen, M., Bentley, S.D., Besra, G.S., Churcher, C., James, K.D., De
Zoraa, A., Chillingworth, T., Cronin, A., Dowd, L., Feltwell, T.,
Hamlin, N., Holtroyd, S., Jagers, K., Moule, S., Quail, M.A.,

```

```

TITLE
JOURNAL Nucleic Acids Res. 31 (22), 6516-6523 (2003)
PUBMED 14602910
REFERENCE 2 (bases 1 to 349535)
AUTHORS Cerdano-Tarraga, A.M.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-2003) Cerdano-Tarraga A.M., submitted on behalf
of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust
Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:
amct@sanger.ac.uk

FEATURES
Location/Qualifiers
1..349535
/organism="Corynebacterium diphtheriae"
/mol_type="genomic DNA"
/strain="NCTC13129"
/db_xref="taxon:1717"
/notes="biotype gravis"
76..402
/locus_tag="DIP1058"
76..402
/locus_tag="DIP1058"
/notes="Similar to Mycobacterium tuberculosis hypothetical
11.0 kDa protein Rv1898 or MT1949 or MTCY180.20C
SM:Y198 MYCTU (007734) (102 aa) fasta scores: E():
5.8e-16 56.43% id in 101 aa, and to Aquifex aeolicus
hypothetical 17.9 kDa protein AQ.2067 TR:067847
(EMBL:AE000771) (157 aa) fasta scores: E(): 4.7e-08,
40.19% id in 102 aa"
401..402
/codon_start=1
/transl_table=11
/product="Conserved hypothetical protein"
/protein_id="CAE49578.1"
/db_xref="GI:38199913"
/translation="MLDMLAFSVAPPTETPNKAEADVAEAEIRVRESGLPNETN
AMFTLGESEVMAVVKATDAVLAVSPRASIVIKADIRPGTGLQOKVAVSERVL
AGDSSES"
88..372
/locus_tag="DIP1058"
/notes="HMPfam hit to PF01910, Protein of unknown function
DUF77"
complement (405..4762)
/notes="Putative iron uptake system. Low GC content
(58.02%)"
complement (405..1226)
/locus_tag="DIP1059"
complement (405..1226)
/locus_tag="DIP1059"
/notes="Similar to Streptomyces coelicolor putative
iron-siderophore uptake system ATP-binding component
SC151.25C TR:09S215 (EMBL:AL109848) (301 aa) fasta scores:
E(): 1.9e-42, 50.58% id in 255 aa, and to Escherichia coli
ferric enterobactin transport ATP-binding protein fepC
b0588 SW:FEPC_ECOLI (P23878) (271 aa) fasta scores: E():
1.1e-41, 49.61% id in 262 aa"
1..402
/transl_table=11
/product="Putative iron-siderophore uptake system
ATP-binding component"
/db_xref="GI:38199914"
/translation="MTSTSPILRDSVTLAWDKHTVSTDLTVEVFGQFTALGPGNC
GKSTLLKSPARILQPHSTQVILGNDPLSLATKYIAQRALLPOTAPADITVEELV
GRGFVYOTLLRQNSTKDAQVDSALLATGTQLRHRRVTELSGQRQVWLAIVLAQ
NGVYVLDSPYTLVLAHQYQLLELALRQQLGRTITVTLDLQAVRYADHLIVMK
GAVYAGAPSSIMTPQLISEVDFINVDHIDRVIVKTPQDPVSVLENLQ"
complement (564..1133)
/locus_tag="DIP1059"
/notes="HMMSmart hit to SM00382, ATPases associated with a
variety of cellular activities"
complement (567..1130)
misc_feature
misc_feature
misc_feature
misc_feature

```

Rabinowitsch, E., Rutherford, K., Thomson, N.R., Unwin, L.,  
Whitehead, S. and Barrell, B.G. Parkhill, J.  
The complete genome sequence and analysis of *Corynebacterium*  
*diphtheriae* NCTC13129  
Nucleic Acids Res. 31 (22), 6516-6523 (2003)  
14602910  
2 (bases 1 to 349535)  
Cerdano-Tarraga, A.M.  
Direct Submission  
Submitted (03-OCT-2003) Cerdano-Tarraga A.M., submitted on behalf  
of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust  
Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:  
amct@sanger.ac.uk

misc\_feature

/locus\_tag="DIP1059"  
/note="HMPfam hit to PF00005, ABC transporter"  
complement(579..797)

misc\_feature

/locus\_tag="DIP1059"  
/note="ProfileScan hit to P850100, 2nd half motif for nucleotide binding, associated with P-loop."  
complement(753..797)  
/locus\_tag="DIP1059"  
/note="ScanRegExp hit to PS00211, ABC transporters family signature."  
complement(1047..1124)

misc\_feature

/locus\_tag="DIP1059"

misc\_feature

/note="ProfileScan hit to P850101, P-loop nucleotide binding motif (does not find all)."  
complement(1086..1109)

gene

/locus\_tag="DIP1059"  
/note="ScanRegExp hit to PS00017, ATP/GTP-binding site motif A (P-loop)."  
complement(1223..2182)  
/locus\_tag="DIP1060"  
complement(1223..2182)

CDS

/locus\_tag="DIP1060"

/note="Similar to Streptomyces coelicolor putative iron-siderophore uptake system transmembrane component sc151.26C TR:Q9S214 (EMBL:AL109848) (375 aa) fasta scores: E(): 5e-40, 40.74% id in 324 aa, and to Escherichia coli ferric enterobactin transport system permease protein fepg fepg or B0589 SW:PEPG\_ECOLI (P23877) (330 aa) fasta scores: E(): 1.8e-45, 44.09% id in 322 aa"  
/codon\_start=1

/transl\_table=11  
/product="Putative iron-siderophore uptake system transmembrane component"  
/protein\_id="CA849582.1"

/db\_xref="GI:38199915"

misc\_feature

/locus\_tag="DIP1060"  
/note="Signal peptide predicted for DIP1060 by SignalP 2.0 HMM (Signal peptide probability 0.966) with cleavage site probability 0.180 between residues 25 and 26;  
signal-peptide site"

misc\_feature

/locus\_tag="DIP1060"

misc\_feature

/note="HMPfam hit to PF01032, FecCD transport family" order(complement(2105..2161), complement(1979..2044), complement(1883..1939), complement(1805..1870), complement(1718..1783), complement(1607..1657), complement(1445..1510), complement(1319..1384), complement(1247..1303))  
/locus\_tag="DIP1060"

misc\_feature

/note="9 probable transmembrane helices predicted for DIP1060 by TMHMM2.0"  
complement(1250..1564)

gene

/locus\_tag="DIP1060"  
/note="BlastProDom hit to PD001557, PD001557"  
complement(2184..3191)

CDS

/locus\_tag="DIP1061"  
complement(2184..3191)

/locus\_tag="DIP1061"  
/note="Similar to Streptomyces coelicolor putative iron-siderophore uptake system transmembrane component sc151.27C TR:Q9S213 (EMBL:AL109848) (348 aa) fasta scores: E(): 8e-36, 36.53% id in 323 aa, and to Escherichia coli ferric enterobactin transport system permease protein fepd or B0590 SW:PEPD\_ECOLI (P23876) (334 aa) fasta scores: E(): 5.2e-40, 40.37% id in 322 aa"  
/codon\_start=1

/transl\_table=11  
/product="Putative iron-siderophore uptake system transmembrane component"  
/protein\_id="CA849584.1"  
/db\_xref="GI:38199916"  
/translation="MLQPRVRLPLVITFLVILLCTIMSLGVGARNITSPDILA AAFYVIGTALAGSLSHISIPALAILGAILAGACVVALSMNSCKSDPLRLVLAGVAL SAILTGIGGLSLVNPQAPDELKSMWGNIDAGSYOPAAVAGVGLILGILVATCMRQ LNALSGDELATMGASIAKTRLFTFLVLAASATAAGVITFLGMVPHIARNIV GPNLLRLVATSLIGLIVLSADILGRIIVPGEPFAGVAVAFIAPFLIAYATKRRK I"  
complement(3090..3191)  
/locus\_tag="DIP1061"  
/note="Signal peptide predicted for DIP1061 by SignalP 2.0 HMM (Signal peptide probability 1.000) with cleavage site probability 0.663 between residues 34 and 35;  
signal-peptide site"  
complement(2199..2525)  
/locus\_tag="DIP1061"  
/note="BlastProDom hit to PD001557, PD001557"  
complement(2205..3095)  
/locus\_tag="DIP1061"  
/note="HMPfam hit to PF01032, FecCD transport family" order(complement(3096..3161), complement(2934..2990), complement(2838..2903), complement(2766..2831), complement(2670..2735), complement(2544..2609), complement(2394..2459), complement(2289..2354), complement(2211..2276))  
/locus\_tag="DIP1061"  
/note="9 probable transmembrane helices predicted for DIP1061 by TMHMM2.0"  
3239..4282  
/locus\_tag="DIP1062"  
3239..4282  
/locus\_tag="DIP1062"  
/note="Similar to Escherichia coli iron(iii) dicitrate-binding periplasmic protein precursor FecB or B4290 SW:PECB\_ECOLI (P15028) (300 aa) fasta scores: E():

Query Match 19.2%; Score 136.6; DB 1; Length 349535;  
Best Local Similarity 52.4%; Pred. No. 2.2e-23;  
Matches 354; Conservative 0; Mismatches 309; Indels 12; Gaps 2;  
QY 2 TGGTGATCATGAATCTTCATTACAGCTCTGCTTTGGGGCCAGTCTTTTACTGTCCA 61  
DB 37321 TGGTAGTATGATTTGCTATGCTGGAATTCCTCATGGGTCTGTCTTGTATCGTGGCA 37262  
QY 62 TCGGACCCGAGATGTACTGGTGAATTAACAAGGAATTAAGCGCGAAGACTCATTTGGG 121  
DB 37261 TCGGGCCACAAAATGCCCTTGATTAATCGTCAAGGCATTAAGCGGAAGGTCTAATCCCCA 37202  
QY 122 TTCTTCTGCTGTTTAAATTTCTGAGCTCTTTTGTTCATCGCGGACCTTTGGGGGTG 181  
DB 37201 TTTGTTGATGATCACTCTCTGAGCTGATTCCTCATTTTGGAGGCACTGCGCGGTGCG 37142  
QY 182 ATCTTTTGTCCAAATCGCGCGCATGCTGCTCGCATATTAATGCGCTGGGTGGCATCGCTT 241  
DB 37141 GCGCATTAAGTCTGATCGAGCCCCATCGCACTAGTAGTACTGAAATGGCTGGTGTAGCGT 37082  
QY 242 ACTGTATGTTGGTTCGTATGCGCAGCGAAGACGCGCATGACAAACAAGGTGGAAGCGC 301  
DB 37081 ACTTGCTCTAATTCGGATTCACTTGTTTTAAAGAAGCATTTAAACGTCAAGCGCAAGCGC 37022  
QY 302 CACAGATCATTTGAAGAACAACCAACCGTGGCCGATCACACGCTTTTGGCGGTTCGG 361  
DB 37021 TCGCATC---GAGCAAGCAAGCAAGCGGTGGCTTACGAAACGATGGCTGACGATCATCG 36965  
QY 362 CGGTGGCCACTGACACCGCAACCGCGGTGCGGTGGAGGTGAGCGTTCGATAAGCAGCGGG 421  
DB 36964 GAGTGATCACTAAACACGCAAAAGCTCAGCCCAAGTCAGC-----GCAGCGCA 36914  
QY 422 TTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGTGTGACCTGTGTGACCCGATTC 481

```

Db      36913  CGTGGGTTAAACAGAGTGTTCACAGCGTGGCTTTACGTGGCTTAATCCAGCTCTTATA 36854
Qy      482   TGGACGCGTTTGTGTTTATCGCGCGGTGGCGCGCAATACGCGCACACCGGACGGTGA 541
Db      36853  TAGACGTTCTTGTGTCATGTTGGAGGATGCCAACACACGCTCCGACGGCGCTGG 36794
Qy      542   TTTTCGCGCGTGGCGCGTTCGCGCAAGCTGATCTGGTTCGCGCTGGTGGGTTTCGGCG 601
Db      36793  TTTTCGCGCGTGGCGCGTTCGCGCAAGCTGATCTGGTTCGCGCTGGTGGGTTTCGGGATACCT 36734
Qy      602   CAGCAGCATTTGTCAGCGCGGTGTCAGCGCCCAAGCTGTCGCGCTGGTCAACGCTCTGCG 661
Db      36733  CCACACGCTTTTACCGTGTATTCGCGCGCAGCGCTGGCGATATCAACATACACAA 36674
Qy      662   TGGCAGTTGTGATGA 676
Db      36673  TCGGAATCATATGA 36659

RESULT 15
LOCUS   AL646059
DEFINITION Ralstonia solanacearum GMI1000 chromosome, complete sequence;
segment 3/19.
ACCESSION AL646059 AL646052
VERSION   AL646059.1 GI:117427391
KEYWORDS  Ralstonia solanacearum
SOURCE    Ralstonia solanacearum
ORGANISM  Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Burkholderiaceae; Ralstonia.
REFERENCE
AUTHORS  Salanoubat,M., Genin,S., Artiguenave,F., Guzy,J., Manganot,S.,
Arlat,M., Billault,A., Brottier,P., Camus,J.C., Cattolico,L.,
Chandler,M., Choigne,N., Claudel-Renard,C., Cunnac,S., Demange,N.,
Gaspin,C., Lavie,M., Moisan,A., Robert,C., Saurin,W., Schiex,T.,
Siquier,P., Thebault,P., Whalen,M., Wincker,P., Levy,M.,
Weissenbach,J. and Boucher,C.A.
Genome sequence of the plant pathogen Ralstonia solanacearum
Nature 415 (6871), 497-502 (2002)
11823852
PUBMED   121691879
REFERENCE 2 (bases 1 to 190050)
Boucher,C.A.
Direct Submission
Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie
Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS,
BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean
Dausset-CEPH, 27 rue Juliette Dodu, 75010 Paris, France, IMGM CNRS
118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA
URGV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,
Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27,
F31326 Castanet-Tolosan Cedex. Laboratoire de Genetique Cellulaire
INRA, BP27, F31326 Castanet-Tolosan Cedex
Christian.Boucher@toulouse.inra.fr
http://sequence.toulouse.inra.fr/R.solanacearum.html.
FEATURES
source    1. .190050
            /organism="Ralstonia solanacearum"
            /mol_type="genomic DNA"
            /strain="GMI1000"
            /db_xref="taxon:305"
            complement(163..750)
            /gene="RSc0382"
            /note="synonym: RS03350"
            complement(163..750)
            /gene="RSc0382"
            /note="synonym: RS03350"
            /function="miscellaneous; not classified regulator"
            /note="Product confidence : probable
            Gene name confidence : hypothetical
            predicted by Codon usage"

gene
CDS
            /gene="RSc0385"
            /function="miscellaneous; hypothetical/global homology"
            /note="Product confidence : probable
            Gene name confidence : hypothetical
            predicted by Codon usage"

```

```

predicted by Homology
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="PROBABLE RESPONSE REGULATOR TRANSCRIPTION
REGULATOR PROTEIN"
/protein_id="CAD13910.1"
/db_xref="GI:17427392"
/db_xref="GOA:Q8Y2F4"
/db_xref="SPTREMBL:Q8Y2F4"
/translation="MAPTVLLVROPDLAEQVNLLEIRAGLAEGFVDVGVVDADLML
PERIAQTQDMVIVASERAAEDTIEHVCVATQHAAPRIVLFTDNDASRIKAAFAAGI
TAIVYGLKPRVXAVLDVAYAPFHEALSAELDNAKTQLAERKVLBRKGLLMQOM
QLSEDAFRRLRKWMDNRNIKLVRAQRVIDVMAG"
937..1638
/gene="RSc0383"
/note="synonym: RS03351"
937..1638
/gene="RSc0383"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD13911.1"
/db_xref="GI:17427393"
/db_xref="GOA:Q8Y2F3"
/db_xref="SPTREMBL:Q8Y2F3"
/translation="MTDFVVTVPVAVGIPVGAAGEPPVRRVVCVRVYAAHAEKMS
DPREPFPFKPADAVRYVADGTGQFYPTETRDCHYEILWALGTGGRIAVES
AAOHIVGYAIGIDMTDRDLQNAKGGKRWETKAFDGSAPIGTIPVPAQTVASPDRA
ITLSVNDKAHQHGLDLSLINSVPETIAYLSKLFELRGDLITGTGPEGVPSIGDIM
VGRIDGVGSLVKVY"
1642..2292
/gene="RSc0384"
/note="synonym: RS03352"
1642..2292
/gene="RSc0384"
/EC_number="5.2.1.2"
/function="small molecule metabolism; degradation; amino
acids"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="PROBABLE MALEYLACETOACETATE ISOMERASE
(GLUTATHIONE TRANSFERASE ZETA 1) PROTEIN"
/protein_id="CAD13912.1"
/db_xref="GI:17427394"
/db_xref="GOA:Q8Y2F2"
/db_xref="SPTREMBL:Q8Y2F2"
/translation="MSIKLYNFRSSAFVRVIALEIKGLPYDVAPVHLLKGEQAAPD
FVKINPEALVPVLCGTEVLNQSLAIVEYLETHPTLPGSAAADRAHIAIALAIA
CTEHLNPRVLKYLKTEVNGDNDNRHWRVIRGFALETRLTQSPRTGACCVGD
TPTLADLCILYQVFNKGRPDVAVEDYPTLARVFHECHWAQPAQRAAPPAQDAASA"
2289..2927
/gene="RSc0385"
/note="synonym: RS03353"
2289..2927
/gene="RSc0385"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon usage"

```

```
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE TRANSMEMBRANE PROTEIN"
/protein_id="CAD13913.1"
/db_xref="GI:17427395"
/db_xref="GOA:O8Y2F1"
/db_xref="SPTREMBL:O8Y2F1"
/translation="MTGPGESAPLIGALSGFLGASLIIVAGIAQNAIVLRQLRRE
YVGLVLIICALCDMALIALGAVAGMTLII SAHPAMLTATVRWAGAFLLAYGARAPRAAW
RGAERLQANGKSHAOVLASALASALILNPHVYLDVTLGALGGRVAMPANVAFAG
GAMCSTLNFSLILGARGALLEPVPFARPVNRVLDALIGAVMAIALTLILMG"
complement(2924..3910)
/gene="ftsY"
/notes="synonyms: RS03354, RS0386"
complement(2924..3910)
/gene="ftsY"
/function="cell processes: cell division"
/notes="Product confidence : putative
Gene name confidence : putative
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE CELL DIVISION PROTEIN"
/protein_id="CAD13914.1"
/db_xref="GI:17427396"
/db_xref="GOA:O8Y2F0"
/db_xref="SPTREMBL:O8Y2F0"
/translation="NPDTPAAPALDMQADDIEVTPTFPVFEQARKGNMHLRSLGSLK
TSKSLTIFVGVDEALFBELETALMADAGVDATEIYLDLDELRRVKAQRIETAVGGV
KTAURLLELLHLEKTMVLGRDQPTVIMTAGVNGAGKTTISGLKFKHFTQYQSVL
LAAGDTFPAARAEQIVWGNRNVTVAQESDPAAVIFDVAANAARGIDIVNADTA
GRLPOLHMEELKKVRVIGKAMATAPHEITLLVIDANTGONALAQVKAFFDALGLTG
LIVTKLDGTAKGUILAARQRPVPVYFIGVGEQVELDPFSAREFADALLG"
4124..5629
/gene="RSC0387"
/notes="synonym: RS03355"
4124..5629
/gene="RSC0387"
/EC number="3.4.99.-"
/function="miscellaneous; hypothetical/global homology"
/notes="Product confidence : putative
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE ZINC PROTEASE PROTEIN"
/protein_id="CAD13915.1"
/db_xref="GI:17427397"
/db_xref="GOA:O8Y2E9"
/db_xref="SPTREMBL:O8Y2E9"
/translation="WRSPQSVNSARTGLRWAGITLAVALLAGEAGAPAPMAARQ
AAALTGAKVGTQADTHEYRLANGLLEIVKEDHRAPTVAHMVYHAGSIDENHGTGTV
AHMLHEMTKTRAVTGPCEFSRVAALGGRENAMTRDFMTYFOQIEKSHLADVMALE
ADRMALQLTKEFPEKMNVKEERMRIDDSARATVYEQMLAVLNAAPVYRPTIGW
PSLDLMTVQDAQDWYHKVAPNNATVVTVGVNPDVEFQAQFTYKLOPHALPRY
AQDEPKVGKIKIWKAPENPVVLYAYKAPPLRDVEKVDVPALEVLISAVLDGIDNA
RLPMLLVKDGDEKGRADDVNGAGIDMNRGFSIFLDDGVDPADGHTTAEIEQALRAQI
DRIARGDVTAEELKRVAAQVVAQIYKRDVSFGQMEIGMAEMTGLSWRDLDRILEKI
KSVTPAQVQVQVAKTYTFTEDNLVATLLPQIDPNKPAKVPQPMREEGGLR"
5629..6972
/gene="RSC0388"
/notes="synonym: RS03356"
5629..6972
```

```
/gene="RSC0388"
/EC number="3.4.99.-"
/function="miscellaneous; hypothetical/global homology"

Query Match      11.8%; Score 84.2; DB 1; Length 190050;
Best Local Similarity 47.9%; Pred. No. 3.9e-10;
Matches 242; Conservative 0; Mismatches 263; Indels 0; Gaps 0;

QY 207 CGTGTCTGATATTATGCGTGGGTGGGATCGCTTACCTGTTATGCTTTGCGGTCATGCG 266
    |||||
Db 2423 CGTGTCTGCGGTGGTCTGATCTTGGCGCTGTGCGCATGCGGCTGATCGGCTGGCGCT 2482
    |||||
QY 267 AGCGAAGACGCCATGATGAAACAAAGCTGGAAGCCACAGATCATTTGAAGAAACAGAAC 326
    |||||
Db 2483 GCGGGCATGGGACCGCTGATCTTGGCGCACCCCGGCATGCTGACCGCGGTGGCGCTGGGC 2542
    |||||
QY 327 AACCGTGGCCGATGACACGCGCTTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCAACCG 386
    |||||
Db 2543 CGCGCGGCGCTTCTCTGCTGGCTTACGCGCACGCGGCTTCCGCGCGGCTGGCGCGCGCG 2602
    |||||
QY 387 GGTGCGGGTGGAGGTGAGCGTGGATACGACGCGGTTTGGGTAAAGCCCATGTTGATGCG 446
    |||||
Db 2603 CGAGCGGCTGACGCGCGGCAACGCGGCACAGGCGCTCGCATGCGCAGGTGCTGGCGCTCCG 2662
    |||||
QY 447 AATCGTGTGACCTGTGAAACCGAATGCGTATTTGGACGCGTTTGTGTTATCGCGCG 506
    |||||
Db 2663 GCTGGCGCTGTGCGTGTCTCAATCCGCACTCTATCTGACACCGTGTGCTGGCGCGC 2722
    |||||
QY 507 GTCGCGCGCGCAATACGCGGACACCGGACCGGTGGAATTTTCGCGCTGGCGGCTTCGCGCG 566
    |||||
Db 2723 CATCGCGCGCGCTTACGCGATGCGCGCAACGTTGGGCTTCCGCGCGCGCGCATGTGCGCG 2782
    |||||
QY 567 AAGCTGTGATCTGTTCCCGCTGGTGGGTTTGGCGGCGCAGCATTTGTACGCCCGCTGTC 626
    |||||
Db 2783 GTGATCTGTTCTTCTGCTGCTGGGTTTGGCGCACCGGCTGTGAGACCGGTGTTGCG 2842
    |||||
QY 627 CAGCCCGCAAGGTGTGGCGCTGGATCAACGTCGTGGTGGCAGTTGTGATGACCGCATTTGCG 686
    |||||
Db 2843 CAGCGCGGTTCGCTGGCGGCTGCTCGATGCGCTGATCGCGCGGCTGATGTGGCGCATCGC 2902
    |||||
QY 687 CATCAACCTGATGTTGATGGGTTAG 711
    |||||
Db 2903 GCTGACGCTGCTGATGGCGGCTTAG 2927
    |||||
```

Search completed: March 15, 2004, 22:02:01  
Job time : 2852.7 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:09:44 ; Search time 289.925 Seconds  
(without alignments)  
10418.122 Million cell updates/sec

Title: US-09-105-117K-1\_COPY\_1016\_1726

Perfect score: 711

Sequence: 1 atgggtgatcgaatctt.....aactgatgttgatgggttag 711

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002s.\*
- 7: Geneseq2003as.\*
- 8: Geneseq2003bs.\*
- 9: Geneseq2003cs.\*
- 10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	711	100.0	711	8	ACC80941
2	711	100.0	822	4	AAF71779
3	711	100.0	822	4	AAS96098
4	711	100.0	2374	2	AAT96816
5	711	100.0	2374	9	ADB66196
6	711	100.0	34980	5	AAB68528
7	708	99.6	708	5	AAB68420
8	696.8	98.0	712	8	ACC80942
9	317.6	44.7	1568	4	AAH45375
10	60.4	8.5	636	3	AAH452691
11	57	8.0	15239	2	AAT33536
12	57	8.0	110000	4	AAI99682_22
13	57	8.0	110000	4	AAI99683_22
14	56.2	7.9	86114	6	ABX09143
15	56.2	7.9	110000	4	AAI99682_05
16	56.2	7.9	110000	4	AAI99683_05
17	47.4	6.7	7521	5	AAH71378
18	47.4	6.7	7521	5	AAH71378
19	46.6	6.6	2000	7	ADA71938
20	44	6.2	2000	7	ADA71938
21	40	5.6	1173	7	ADA71370
22	39.4	5.5	624	7	ACF71727
23	39.4	5.5	110000	7	ACF67367_49

24	39.4	5.5	110000	7	ACF65387_0
25	39	5.5	993	4	AAF71777
26	39	5.5	993	4	AAS96096
27	39	5.5	2823	7	ACA38373
28	39	5.5	2826	7	ACA40558
29	39	5.5	110000	4	AAI99682_20
30	39	5.5	110000	4	AAI99683_20
31	38.6	5.4	1245	7	ACA26476
32	37.8	5.3	536	9	ADB68842
33	37.6	5.3	536	9	ADB68842
34	37.6	5.3	1149	7	ACF39391
35	37.6	5.3	110000	4	AAI99682_07
36	37.6	5.3	110000	4	AAI99683_07
37	37.4	5.3	1350	7	ACA27310
38	37.2	5.2	424	7	ABX47227
39	37.2	5.2	873	7	ACA37840
40	37.2	5.2	1509	7	ABX56070
41	37	5.2	945	7	ACA26592
42	37	5.2	1155	5	AAF26425
43	37	5.2	1464	7	ACA43878
44	36.8	5.2	678	7	ACA38495
45	36.8	5.2	681	7	ACA40313

## ALIGNMENTS

RESULT 1					
ACC80941	ID	ACC80941 standard; DNA; 711 BP.			
XX	AC	ACC80941;			
XX	DT	27-OCT-2003 (revised)			
DT	11-AUG-2003	(first entry)			
XX	DE	LyseE protein encoding sequence.			
XX	KW	L-lysine; L-arginine; LyseE; ds.			
XX	OS	Corynebacterium glutamicum.			
XX	FH	Key Location/Qualifiers			
FT	CDS	1..711			
FT		/*tag= a			
FT		/product= "lyseE protein"			
XX	PN	EP1266966-A2.			
XX	PD	18-DEC-2002.			
XX	PF	05-JUN-2002; 2002EP-00012539.			
XX	PR	12-JUN-2001; 2001JP-00177075.			
XX	PA	(AJIN ) AJINOMOTO CO INC.			
XX	PI	Gunji Y, Yasueda H;			
XX	DR	WPI; 2003-241171/24.			
XX	DR	P-PSDB; ABR58213.			
XX	PT	Novel DNA encoding variant of LyseE protein from a coryneform bacterium, when introduced into methanol assimilating bacterium, facilitates excretion of L-lysine and/or L-arginine to outside of a cell.			
XX	PS	Example 1; Page 17-18; 23pp; English.			
XX	CC	The present invention relates to DNA encoding variants of protein with loop region and six hydrophobic helices which facilitates excretion of L-lysine and/or L-arginine to outside of cell of a methanol assimilating bacterium when introduced into the bacterium. The method is used for encoding a protein which facilitates excretion of L-lysine, L-arginine or			



CC both of these L-amino acids to outside of a cell of a methanol  
CC assimilating bacterium when DNA of the method is introduced into the  
CC bacterium. The present sequence represents a lysin protein from  
CC Brevibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003  
CC to standardise OS field)

XX  
SQ Sequence 711 BP; 135 A; 173 C; 222 G; 181 T; 0 U; 0 Other;

Query Match 100.0%; Score 711; DB 8; Length 711;  
Best Local Similarity 100.0%; Pred. No. 3.9e-206;

Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCATCATCGAATCTTCATTACAGTCTGCTTTTGGGCGCCAGCTTTTACTGTCC 60  
Db |||||  
QY 1 ATGTCATCATCGAATCTTCATTACAGTCTGCTTTTGGGCGCCAGCTTTTACTGTCC 60  
Db |||||  
QY 61 ATCGGACCGCAGATGTAAGTAAACAAAGGAATTAAGCGCGAAGGACTCAATTGGC 120  
Db |||||  
QY 121 GTTCTTCGTGTTTAAATTCAGAGTCTTTTGTTCATCGCGCACCTTGGGGCTT 180  
Db |||||  
QY 181 GATCTTTTGTCAATGCGCGCGATCGTCTCGATATTATGGCTGGGGTGGCATCGCT 240  
Db |||||  
QY 241 TACCTGTTATGTTTGGCGTATGGCAGCGAAAGACCGCATGACAAACAGGTGGAGCG 300  
Db |||||  
QY 301 CCACAGATCATTTGAAGAACAGACCAACCGGTCGCCGATGACACGCTTTGGCGGTTCC 360  
Db |||||  
QY 361 GCGTGGCCACTGACACCGCGCAACCGGTCGCCGATGAGCGTGGAGTGGATAGCAAGCAGCG 420  
Db |||||  
QY 421 GTTGGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGGTTGAACCGCAATCGCTAT 480  
Db |||||  
QY 481 TTGACCGGTTGCTTTTATCGCGCGTGGCGGCAATACCGCGACACCGCGAGTGG 540  
Db |||||  
QY 541 ATTTTCGCCGCTGGCGCGTTCGCGCAAGCCTGATCTGTTTCCCGCTGGTGGTTTCGC 600  
Db |||||  
QY 601 GCAGCAGATTGTCAGCGCGCTGTCAGCGCCCAAGGTGGCGTGGATCAAGTCGTC 660  
Db |||||  
QY 661 GTTGCAGTTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGTTAG 711  
Db |||||

RESULT 2

AAAF71779  
ID AAF71779 standard; DNA; 822 BP.  
XX AAF71779;  
AC AAF71779;  
XX  
XX  
DT 30-APR-2001 (first entry)  
DE Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:53.  
XX Corynebacterium glutamicum; metabolic pathway protein; MP protein;  
KW fine chemical production; microorganism; organic acid; nucleoside;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;  
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.  
OS  
XX Corynebacterium glutamicum.  
XX WO200100843-A2.  
XX 04-JAN-2001.  
XX 23-JUN-2000; 2000WO-IB000923.  
XX 25-JUN-1999; 99US-0141031P.  
PR 01-JUL-1999; 99DE-01030476.  
PR 02-JUL-1999; 99US-0142101P.  
PR 08-JUL-1999; 99DE-01031415.  
PR 08-JUL-1999; 99DE-01031418.  
PR 08-JUL-1999; 99DE-01031419.  
PR 08-JUL-1999; 99DE-01031420.  
PR 08-JUL-1999; 99DE-01031424.  
PR 08-JUL-1999; 99DE-01031428.  
PR 08-JUL-1999; 99DE-01031434.  
PR 08-JUL-1999; 99DE-01031435.  
PR 08-JUL-1999; 99DE-01031443.  
PR 08-JUL-1999; 99DE-01031453.  
PR 08-JUL-1999; 99DE-01031457.  
PR 08-JUL-1999; 99DE-01031465.  
PR 08-JUL-1999; 99DE-01031478.  
PR 08-JUL-1999; 99DE-01031510.  
PR 08-JUL-1999; 99DE-01031541.  
PR 08-JUL-1999; 99DE-01031573.  
PR 08-JUL-1999; 99DE-01031592.  
PR 08-JUL-1999; 99DE-01031632.  
PR 08-JUL-1999; 99DE-01031634.  
PR 08-JUL-1999; 99DE-01031636.  
PR 09-JUL-1999; 99DE-01032125.  
PR 09-JUL-1999; 99DE-01032136.  
PR 09-JUL-1999; 99DE-01032186.  
PR 09-JUL-1999; 99DE-01032206.  
PR 09-JUL-1999; 99DE-01032227.  
PR 09-JUL-1999; 99DE-01032228.  
PR 09-JUL-1999; 99DE-01032229.  
PR 09-JUL-1999; 99DE-01032230.  
PR 14-JUL-1999; 99DE-01032922.  
PR 14-JUL-1999; 99DE-01032926.  
PR 14-JUL-1999; 99DE-01032928.  
PR 14-JUL-1999; 99DE-01033004.  
PR 14-JUL-1999; 99DE-01033005.  
PR 14-JUL-1999; 99DE-01033006.  
PR 12-AUG-1999; 99US-0148613P.  
PR 27-AUG-1999; 99DE-01040764.  
PR 27-AUG-1999; 99DE-01040765.  
PR 27-AUG-1999; 99DE-01040766.  
PR 27-AUG-1999; 99DE-01040832.  
PR 31-AUG-1999; 99DE-01041378.  
PR 31-AUG-1999; 99DE-01041379.  
PR 31-AUG-1999; 99DE-01041380.  
PR 31-AUG-1999; 99DE-01041394.  
PR 31-AUG-1999; 99DE-01041396.  
PR 03-SEP-1999; 99DE-01042076.  
PR 03-SEP-1999; 99DE-01042077.  
PR 03-SEP-1999; 99DE-01042079.  
PR 03-SEP-1999; 99DE-01042086.  
PR 03-SEP-1999; 99DE-01042087.  
PR 03-SEP-1999; 99DE-01042088.  
PR 03-SEP-1999; 99DE-01042095.  
PR 03-SEP-1999; 99DE-01042124.  
PR 03-SEP-1999; 99DE-01042129.  
PR 09-MAR-2000; 2000US-0187970P.  
XX (BADI ) BASF AG.  
XX

752 GTGGCAGTTGTGATGCA CCGCATTTGGCCATCAACATGATGTTGATGGGTTAG B02

db 272 GATCTTTGTCCAAATGCCGGGCGGATCGTGTTCGATTTATGGCGTGGGGCATCGCT 331

QY 241 TACCTGTTATGTTTGGCGTCATGCGAGCAAGAACGCCATGACAAACAGGTTGGAAGCG 300  
 Db 332 TACCTGTTATGTTTGGCGTCATGCGAGCAAGAACGCCATGACAAACAGGTTGGAAGCG 391  
 QY 301 CCACAGATCATTTGAGAAACAGAACCAACCGTCCGATGACACACGCTTTGGGCGGTTGCG 360  
 Db 332 CCACAGATCATTTGAGAAACAGAACCAACCGTCCGATGACACACGCTTTGGGCGGTTGCG 451  
 QY 361 GCGGTGGCCACTGACACACGCGCAACCGGTTGGGTTGAGGTTGAGGTCGATGACGAGCGG 420  
 Db 452 GCGGTGGCCACTGACACACGCGCAACCGGTTGGGTTGAGGTTGAGGTCGATGACGAGCGG 511  
 QY 421 GTTTGGGTAAAGCCCATGTTGATGCGCAATCTGCTGACCTGTTGAACCCCAATGCGTAT 480  
 Db 512 GTTTGGGTAAAGCCCATGTTGATGCGCAATCTGCTGACCTGTTGAACCCCAATGCGTAT 571  
 QY 481 TTGGAGCGGTTTGTGTTATCGGCGGCTGCGCGGCAATACGCGACACCGGACGCGTGG 540  
 Db 572 TTGGAGCGGTTTGTGTTATCGGCGGCTGCGCGGCAATACGCGACACCGGACGCGTGG 631  
 QY 541 ATTTTTCGCGCTGGCGGCTTCGCGCAAGCCTGATCTGTTCCCGCTGCTGGTTTCGCG 600  
 Db 632 ATTTTTCGCGCTGGCGGCTTCGCGCAAGCCTGATCTGTTCCCGCTGCTGGTTTCGCG 691  
 QY 601 GCACGACATTTGACGCGGCTGTCAGCGCCAGGTTGCGGCTGCGGATCAAGTCGTC 660  
 Db 692 GCACGACATTTGACGCGGCTGTCAGCGCCAGGTTGCGGCTGCGGATCAAGTCGTC 751  
 QY 661 GTGGCACTTGTGATGACCGCAATTTGGCCATCAAACTGATGTTGATGGGTTAG 711  
 Db 752 GTGGCACTTGTGATGACCGCAATTTGGCCATCAAACTGATGTTGATGGGTTAG 802

RESULT 4  
 AAT96816  
 ID AAT96816 standard; DNA; 2374 BP.  
 XX AAT96816;  
 XX 12-MAR-1998 (first entry)  
 XX DNA encoding LysG, LysE and ORF3 from *Corynebacterium glutamicum*.  
 XX LysG; LysE; ORF3; lysine transport; regulatory protein; export protein;  
 KW Microbial production; amino acid; animal feed additive; ds.  
 XX *Corynebacterium glutamicum*.  
 FH Key Location/Qualifiers  
 CDS complement(82..954)  
 FT /\*tag= a  
 FT /label= LysG  
 FT 1016..1726  
 FT /\*tag= b  
 FT /label= LysE  
 FT complement(1723..2373)  
 FT /\*tag= c  
 FT /label= orf3  
 XX DE19548222-A1.  
 XX 26-JUN-1997.  
 XX 22-DEC-1995; 95DE-01048222.  
 XX 22-DEC-1995; 95DE-01048222.  
 XX (KERJ ) FORSCHUNGSZENTRUM JUELICH GMBH.  
 XX Vrijic M, Eggeling L, Sahm H;  
 XX WPI; 1997-333867/31.

DR P-PSDB; AAW37714, AAW37715, AAW37716.  
 XX Increasing microbial production of amino acids, especially lysine - by  
 PT improving export carrier activity or corresponding gene expression, also  
 PT new export and regulatory genes from *Corynebacterium*.  
 XX Claim 23 and 26; Page; 16pp; German.  
 XX This DNA, isolated from *Corynebacterium glutamicum*, contains the LysG,  
 CC LysE and ORF3 genes. LysG and LysE encode a lysine transport regulatory  
 CC protein and an export protein, respectively. Microbial production of  
 CC amino acids (A) is improved by increasing the export-carrier activity  
 CC and/or the export gene expression in a microorganism that produces (A).  
 CC The method is specifically used to increase production of lysine, used as  
 CC an animal feed additive. Other (A) are variously useful as  
 CC pharmaceuticals, condiments and intermediates for fine chemicals. This  
 CC method increases the amount of (A) secreted into the culture medium.  
 CC Export of (A) has been found to depend on a single gene. NB. This  
 CC sequence has been created from the information given in table 2 of the  
 CC specification  
 XX Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;  
 SQ Query Match 100.0%; Score 711; DB 2; Length 2374;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-206;  
 Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGTGATCATGGAATCTTCATTACAGGTCGCTTTTGGGGGCCAGTCTTTTACTGTCC 60  
 Db 1016 ATGGTGATCATGGAATCTTCATTACAGGTCGCTTTTGGGGGCCAGTCTTTTACTGTCC 1075  
 QY 61 ATCGGACCGCAGAAATGTAAGCAAGAAATTAAGCGCGAAGGACTCATTGCG 120  
 Db 1076 ATCGGACCGCAGAAATGTAAGCAAGAAATTAAGCGCGAAGGACTCATTGCG 1135  
 QY 121 GTTCTTCTCGTGTTTAAATTTCTGAGCTCTTTTGTTCATCGCGCGGACCTTGGCGTT 180  
 Db 1136 GTTCTTCTCGTGTTTAAATTTCTGAGCTCTTTTGTTCATCGCGCGGACCTTGGCGTT 1195  
 QY 181 GATCTTTTGTCCAATCGCGCGCATCTGCTCGATATTATGCTGTGGGTTGGCATCGCT 240  
 Db 1196 GATCTTTTGTCCAATCGCGCGCATCTGCTCGATATTATGCTGTGGGTTGGCATCGCT 1255  
 QY 241 TACCTGTTATGTTTCCGCTCATGCGCAGCAAGAGCGCCATGACAAACAGGTTGGAAGCG 300  
 Db 1256 TACCTGTTATGTTTCCGCTCATGCGCAGCAAGAGCGCCATGACAAACAGGTTGGAAGCG 1315  
 QY 301 CCACAGATCATTTGAGAAACAGAACCAACCGTCCGATGACACACGCTTTGGGCGGTTGCG 360  
 Db 1316 CCACAGATCATTTGAGAAACAGAACCAACCGTCCGATGACACACGCTTTGGGCGGTTGCG 1375  
 QY 361 GCGGTGGCCACTGACACACGCGCAACCGGTTGGGTTGAGGTTGAGGTCGATGAGCGG 420  
 Db 1376 GCGGTGGCCACTGACACACGCGCAACCGGTTGGGTTGAGGTTGAGGTCGATGAGCGG 1435  
 QY 421 GTTTGGGTAAAGCCCATGTTGATGCGCAATCTGCTGACCTGTTGAACCCCAATGCGTAT 480  
 Db 1436 GTTTGGGTAAAGCCCATGTTGATGCGCAATCTGCTGACCTGTTGAACCCCAATGCGTAT 1495  
 QY 481 TTGGAGCGGTTTGTGTTATCGGCGGCTGCGCGGCAATACGCGACACCGGACGCGTGG 540  
 Db 1496 TTGGAGCGGTTTGTGTTATCGGCGGCTGCGCGGCAATACGCGACACCGGACGCGTGG 1555  
 QY 541 ATTTTTCGCGCTGGCGGCTTCGCGCAAGCCTGATCTGTTCCCGCTGCTGGTTTCGCG 600  
 Db 1556 ATTTTTCGCGCTGGCGGCTTCGCGCAAGCCTGATCTGTTCCCGCTGCTGGTTTCGCG 1615  
 QY 601 GCACGACATTTGACGCGGCTGTCAGCGCCAGGTTGCGGCTGCGATCAAGTCGTC 660  
 Db 1616 GCACGACATTTGACGCGGCTGTCAGCGCCAGGTTGCGGCTGCGATCAAGTCGTC 1675  
 QY 661 GTGGCACTTGTGATGACCGCAATTTGGCCATCAAACTGATGTTGATGGGTTAG 711

Db 1676 GTGGCAGTTGTGATGACCGGATTCGCGCATCAACTGATCTGTGATGGGTTAG 1726

RESULT 5  
ADB66196  
ID ADB66196 standard; DNA; 2374 BP.  
XX AC ADB66196;  
XX DT 04-DEC-2003 (first entry)  
XX DE DNA fragment containing C. glutamicum lysG and lysE genes.  
XX KW L-arginine production; coryneform bacteria; lysE, arginine repressor;  
XX KW argE; liver function promoting agent; amino acid infusion;  
XX KW amino acid pharmaceutical; lysG; ds.  
XX OS Corynebacterium glutamicum.  
XX EH Key Location/Qualifiers  
XX FT CDS 1025..1726  
XX FT /\*tag= a  
XX FT /product= "Protein encoded by lysE gene"  
XX PN US2003113899-A1.  
XX PS 19-JUN-2003.  
XX PF 17-JUL-2002; 2002US-00196232.  
XX PR 25-JUL-2001; 2001JP-00224586.  
XX PA (AJIN ) AJINOMOTO CO INC.  
XX PI Yanaguchi M, Ito H, Gunji Y, Yasueda H;  
XX DR WPI; 2003-708853/67.  
XX DR P-PSDB; ADB66197.  
XX

A microorganism comprising enhanced expression of the lysE gene is useful for enhanced production of L-arginine.

Example 4; Page 26-27; 36pp; English.

The present invention relates to a method for producing L-arginine in a microorganism (e.g. coryneform bacteria) that has L-arginine producing ability and has been modified for enhanced expression of the lysE gene. The microorganism is also modified so that an arginine repressor (argR) does not function normally. The method of the invention is useful for the enhanced production of L-arginine which is useful in liver function promoting agents, amino acid infusion and comprehensive amino acid pharmaceuticals. The present sequence represents a DNA fragment containing Corynebacterium glutamicum lysG and lysE genes. Note: The present sequence is given as SEQ ID No:24 in the Sequence Listing but is referred to as SEQ ID No:25 in the rest of the specification.

XX Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;  
SQ

Query Match 100.0%; Score 711; DB 9; Length 2374;  
Best Local Similarity 100.0%; Pred. No. 6.9e-206;  
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGATCATGGAATCTTCATTACAGGCTCTGCTTTGGGGCCAGCTCTTTTACTGTCC 60  
Db |||||  
QY 61 ATCGACCGCAGATGTACTGTGATTAAACAAGGAATTAAGCGCAGGACTCATTTGCG 120  
Db |||||  
QY 1076 ATCGACCGCAGATGTACTGTGATTAAACAAGGAATTAAGCGCAGGACTCATTTGCG 1135  
Db |||||  
QY 121 GTTCTTCTGCTGTTTAATTCTGACGCTCTTTTGTTCATCGCGGCACTTGGCGTT 180  
Db |||||  
QY 1136 GTTCTTCTGCTGTTTAATTCTGACGCTCTTTTGTTCATCGCGGCACTTGGCGTT 1195  
Db |||||

QY 181 GATCTTTTGTCCATFCCGCGCGCATCGTGTCTGCATATTATGCGCTGGGGTGGCATCGCT 240  
Db |||||  
QY 1196 GATCTTTTGTCCATFCCGCGCGCATCGTGTCTGCATATTATGCGCTGGGGTGGCATCGCT 1255  
Db |||||  
QY 241 TACCTGTTATGTTTGGCGTCATGCGCAGCAAGACGCCATGACAAACAAGSTGGAGCG 300  
Db |||||  
QY 1256 TACCTGTTATGTTTGGCGTCATGCGCAGCAAGACGCCATGACAAACAAGSTGGAGCG 1315  
Db |||||  
QY 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTGCCTCCATGACACGCCCTTTGGGCGGTTG 360  
Db |||||  
QY 1316 CCACAGATCATTTGAAGAAACAGAACCAACCGTGCCTCCATGACACGCCCTTTGGGCGGTTG 1375  
Db |||||  
QY 361 CGCGTGGCCACTGACACCGCCAAACCGGTCGCGTGGAGTGAGCTGATGAAGCAGCGG 420  
Db |||||  
QY 1376 CGCGTGGCCACTGACACCGCCAAACCGGTCGCGTGGAGTGAGCTGATGAAGCAGCGG 1435  
Db |||||  
QY 421 GTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGGTTGAACCCGAATGCGTAT 480  
Db |||||  
QY 1436 GTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGGTTGAACCCGAATGCGTAT 1495  
Db |||||  
QY 481 TTGACGCGCTTTGTTGTTTATCGCGCGCTGCGCGCAATAGCGGACACCGCGAGCGTGG 540  
Db |||||  
QY 1496 TTGACGCGCTTTGTTGTTTATCGCGCGCTGCGCGCAATAGCGGACACCGCGAGCGTGG 1555  
Db |||||  
QY 541 ATTTTGGCGCTGGCGCGCTTTCGCGGCAAGCTGATCTGCTCCCGTGGTGGTTTGGC 600  
Db |||||  
QY 1556 ATTTTGGCGCTGGCGCGCTTTCGCGGCAAGCTGATCTGCTCCCGTGGTGGTTTGGC 1615  
Db |||||  
QY 601 GCAGCAGCATTTGTCAACGCCGCTGTCCAGCCCCCAAGGTGTGGCGCTGGATCAAGCTGTC 660  
Db |||||  
QY 1616 GCAGCAGCATTTGTCAACGCCGCTGTCCAGCCCCCAAGGTGTGGCGCTGGATCAAGCTGTC 1675  
Db |||||  
QY 661 GTGCGAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTGATGGGTTAG 711  
Db |||||  
QY 1676 GTGCGAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTGATGGGTTAG 1726  
Db |||||

RESULT 6  
AAH68528/c  
ID AAH68528 standard; DNA; 349980 BP.  
XX AC AAH68528;  
XX DT 26-SEP-2001 (first entry)  
XX DE C glutamicum coding sequence fragment SEQ ID NO: 7063.  
XX KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
XX KW organic acid synthesis; ds.  
XX OS Corynebacterium glutamicum.  
XX PN EF1108790-A2.  
XX PD 20-JUN-2001.  
XX PF 18-DEC-2000; 2000EP-00127688.  
XX PR 16-DEC-1999; 99JP-00377484.  
XX PR 07-APR-2000; 2000JP-00159162.  
XX PR 03-AUG-2000; 2000JP-00280988.  
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.  
XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX DR WPI; 2001-376931/40.  
XX

Novel polynucleotides derived from Coryneform bacteria, for identifying mutation point of a gene, measuring expression of a gene, analyzing expression profile or pattern of a gene and identifying homologous gene.

XX PS Disclosure; SEQ ID NO 7063; 246pp + Sequence Listing; English.

XX CC The present invention provides a number of nucleotide and protein

CC sequences from the Corynebacterium glutamicum glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of corynebacterium, measuring expression amount and analysing

CC the expression profile or expression pattern of a gene derived from

CC Corynebacterium, and identifying a homologue of a gene derived from

CC corynebacterium. Corynebacterium bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a nucleic acid described

CC in the exemplification of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from the European Patent Office

XX SQ Sequence 349980 BP; 80727 A; 91049 C; 97618 G; 80586 T; 0 U; 0 Other;

Query Match 100.0%; Score 711; DB 5; Length 349980;

Best Local Similarity 100.0%; Pred. No. 7.2e-205; Indels 0; Gaps 0;

Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGATCATCGAAATCTTCATTACAGCTCTGCTTTTGGGGCCAGCTTTTACTGTCC 60

Db 128953 ATGGTGATCATCGAAATCTTCATTACAGCTCTGCTTTTGGGGCCAGCTTTTACTGTCC 128894

QY 61 ATCGGACCGCAGATGACTGTGATTAACAGGAATTAAGCGCGAAGGACTCATTCGG 120

Db 128893 ATCGGACCGCAGATGACTGTGATTAACAGGAATTAAGCGCGAAGGACTCATTCGG 128834

QY 121 GTTCTCTCGTGTTAAATTTCTGACGCTCTTTTTCATCCCGGACCTTGGGCGTT 180

Db 128833 GTTCTCTCGTGTTAAATTTCTGACGCTCTTTTTCATCCCGGACCTTGGGCGTT 128774

QY 181 GATCTTTTGTCAATGCCCGCGCATCGTCTCGATATTAATGCGTGTGGGTGATCGCT 240

Db 128773 GATCTTTTGTCAATGCCCGCGCATCGTCTCGATATTAATGCGTGTGGGTGATCGCT 128714

QY 241 TACTGTATGTTTGGCGTCATGCGCAGGAGGAGCCGATGACAAACAGTGGAGCG 300

Db 128713 TACTGTATGTTTGGCGTCATGCGCAGGAGGAGCCGATGACAAACAGTGGAGCG 128654

QY 301 CCACAGATCATTTGAGAAACAGAACCAACCGTGCCCGATGACACGCTTTGGGCGTTG 360

Db 128653 CCACAGATCATTTGAGAAACAGAACCAACCGTGCCCGATGACACGCTTTGGGCGTTG 128594

QY 361 GCGGTGGCCACTGACACGCGCAACCGCGGTGCGGTGGAGGTGAGCTGATAGCAGCG 420

Db 128593 GCGGTGGCCACTGACACGCGCAACCGCGGTGCGGTGGAGGTGAGCTGATAGCAGCG 128534

QY 421 GTTTGGGTAAAGCCCATGTTGATGCAATCGTCTGACTGTTGAACCCGAAATGCGTAT 480

Db 128533 GTTTGGGTAAAGCCCATGTTGATGCAATCGTCTGACTGTTGAACCCGAAATGCGTAT 128474

QY 481 TTGACACGCTTTGTTTATCGGGCGGTGCGCGCGCAATACGGCGACACCGGAGTGG 540

Db 128473 TTGACACGCTTTGTTTATCGGGCGGTGCGCGCGCAATACGGCGACACCGGAGTGG 128414

QY 541 ATTTTCGCGCTGGCGGTGCGCGCAAGCTGATCTGGTTCGCTGTGGGTTCGCG 600

Db 128413 ATTTTCGCGCTGGCGGTGCGCGCAAGCTGATCTGGTTCGCTGTGGGTTCGCG 128354

QY 601 GCACGACGATTTGACGCGCGCTGTCAGCCCGGAGGCTGCGCTGATCAACGTCGTC 660

Db 128353 GCACGACGATTTGACGCGCGCTGTCAGCCCGGAGGCTGCGCTGATCAACGTCGTC 128294

QY 661 GTGGCAGTTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAG 711

Db 128293 GTGGCAGTTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAG 128243

ID XX AAH68420 standard; DNA; 708 BP.

AC XX AAH68420;

DT XX 26-SEP-2001 (first entry)

DE XX C glutamicum coding sequence fragment SEQ ID NO: 3455.

KW Corynebacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis; ds.

OS Corynebacterium glutamicum.

XX BP1108790-A2.

PN 20-JUN-2001.

PD XX 19-DEC-2000; 2000EP-00127688.

PF XX 16-DEC-1999; 99JP-00377484.

PR XX 07-APR-2000; 2000JP-00159162.

PR XX 03-AUG-2000; 2000JP-00280988.

XX (KYOW ) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

DR P-PSDB; AAG93201.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying

XX mutation point of a gene, measuring expression of a gene, analyzing

XX expression profile or pattern of a gene and identifying homologous gene.

PS Claim 1; SEQ ID NO 3455; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

CC sequences from the Corynebacterium glutamicum glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of corynebacterium, measuring expression amount and analysing

CC the expression profile or expression pattern of a gene derived from

CC Corynebacterium, and identifying a homologue of a gene derived from

CC corynebacterium. Corynebacterium bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a nucleic acid described

CC in the exemplification of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from the European Patent Office

XX SQ Sequence 708 BP; 134 A; 173 C; 221 G; 180 T; 0 U; 0 Other;

Query Match 99.6%; Score 708; DB 5; Length 708;

Best Local Similarity 100.0%; Pred. No. 3.2e-205; Indels 0; Gaps 0;

Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGATCATCGAAATCTTCATTACAGGCTGCTTTTGGGGCCAGCTTTTACTGTCC 60

Db 1 ATGGTGATCATCGAAATCTTCATTACAGGCTGCTTTTGGGGCCAGCTTTTACTGTCC 60

QY 61 ATCGGACCGCAGATGACTGTGATTAACAGGAATTAAGCGCGAAGGACTCATTCGG 120

Db 61 ATCGGACCGCAGATGACTGTGATTAACAGGAATTAAGCGCGAAGGACTCATTCGG 120

QY 121 GTTCTCTCGTGTTAAATTTCTGACGCTCTTTTTCATCCCGGACCTTGGGCGTT 180

Db 121 GTTCTCTCGTGTTAAATTTCTGACGCTCTTTTTCATCCCGGACCTTGGGCGTT 180

QY 181 GATCTTTTGTCAATGCCCGCGCATCGTCTCGATATTAATGCGTGTGGGTGATCGCT 240

Db 181 GATCTTTTGTCAATGCCCGCGCATCGTCTCGATATTAATGCGTGTGGGTGATCGCT 240

QY 241 TACTGTATGTTTGGCGTCATGCGCAGGAGGAGCCGATGACAAACAGTGGAGCG 300

Db 241 TACCTGATGTTTCCGTCATGCGAGGAGAGCGGATGACAAACAGGTGGAAGCG 300  
Qy 301 CCACAGATCATTTCAAGAAACAGAACACCGTGGCCGATGACACGCTTTTGGCGGTTGG 360  
Db 301 CCACAGATCATTTCAAGAAACAGAACACCGTGGCCGATGACACGCTTTTGGCGGTTGG 360  
Qy 361 GCGGTGGCCACTGACACGCGCAACCGGTTGGGTTGGAGTGGAGTGGATGAAGCGGG 420  
Db 361 GCGGTGGCCACTGACACGCGCAACCGGTTGGGTTGGAGTGGAGTGGATGAAGCGGG 420  
Qy 421 GTTTCGGTAAAGCCCATGTTGATGCAATGCTGCTGACCTGGTTGAACCCGATCGGTAT 480  
Db 421 GTTTCGGTAAAGCCCATGTTGATGCAATGCTGCTGACCTGGTTGAACCCGATCGGTAT 480  
Qy 481 TTGGACGCGTTTGTGTTTATCGCGCGGCTCGCGCGCAATACGCGCAACCGGACGTTGG 540  
Db 481 TTGGACGCGTTTGTGTTTATCGCGCGGCTCGCGCGCAATACGCGCAACCGGACGTTGG 540  
Qy 541 ATTTTCGCGCTGGCGGTTTCGCGCAAGCTGATCTGGTTCCCGCTGGTGGTTTGGC 600  
Db 541 ATTTTCGCGCTGGCGGTTTCGCGCAAGCTGATCTGGTTCCCGCTGGTGGTTTGGC 600  
Qy 601 GCAGCAGCATTTGTCAGCGCGCTGTCAGCGCCCAAGGTTGGCGCTGGATCAACGTCGTC 660  
Db 601 GCAGCAGCATTTGTCAGCGCGCTGTCAGCGCCCAAGGTTGGCGCTGGATCAACGTCGTC 660  
Qy 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGT 708  
Db 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGT 708

RESULT 8

ACC80942  
ID ACC80942 standard; DNA; 712 BP.

XX AC ACC80942;  
XX 27-OCT-2003 (revised)  
DT 11-AUG-2003 (first entry)

XX LysE24 protein encoding sequence.

XX L-lysine, L-arginine; LysE24; da.

XX Corynebacterium glutamicum.

XX Key Location/Qualifiers  
FH 1..711  
FT CDS /\*tag= a  
FT /product= "lysE protein"

XX EP1266966-A2.

XX 18-DEC-2002.

XX 05-JUN-2002; 2002EP-00012539.

XX 12-JUN-2001; 2001JP-00177075.

XX (AJIN ) AJINOMOTO CO INC.

XX Gunji Y, Yasueda H;

XX WPI; 2003-241171/24.

XX P-PSDB; ABR58214.

XX Novel DNA encoding variant of LysE protein from a coryneform bacterium,  
PT when introduced into methanol assimilating bacterium, facilitates  
XX excretion of L-lysine and/or L-arginine to outside of a cell.

XX Example 1; Page 19-20; 23pp; English.

CC The present invention relates to DNA encoding variants of protein with  
CC loop region and six hydrophobic helices which facilitates excretion of L-  
CC lysine and/or L-arginine to outside of cell of a methanol assimilating  
CC bacterium when introduced into the bacterium. The method is used for  
CC encoding a protein which facilitates excretion of L-lysine, L-arginine or  
CC both of these L-amino acids to outside of a cell of a methanol  
CC assimilating bacterium when DNA of the method is introduced into the  
CC bacterium. The present sequence represents a LysE24 protein from  
CC Brevibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003  
CC to standardise OS field)

XX  
SQ Sequence 712 BP; 133 A; 173 C; 224 G; 182 T; 0 U; 0 Other;

Query Match 98.0%; Score 696.8; DB 8; Length 712;  
Best Local Similarity 99.6%; Pred No. 8.3e-202;  
Matches 709; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 ATGGTGATCATGGAATCTTCAATACAGGTCTGCTTTTGGGGCCAGTCTTTTACGTGCC 60  
Db 1 ATGGTGATCATGGAATCTTCAATACAGGTCTGCTTTTGGGGCCAGTCTTTTACGTGCC 60

Qy 61 ATCGACCGCGAATGTACTGTGATTAAACAGGAATTAAGCGCGAGGACTCATTTGG 120  
Db 61 ATCGACCGCGAATGTACTGTGATTAAACAGGAATTAAGCGCGAGGACTCATTTGG 120

Qy 121 GTTCTTCTGCTGTTTAAATTTCTGACGCTCTTTTGTTCATCGCGGCACTTTGGGCGTT 180  
Db 121 GTTCTTCTGCTGTTTAAATTTCTGACGCTCTTTTGTTCATCGCGGCACTTTGGGCGTT 180

Qy 181 GATCTTTTGTCAATGCGCGCGGATCTGCTGCTGATTAATGCGCTGGGGTGCATCGGT 240  
Db 181 GATCTTTTGTCAATGCGCGCGGATCTGCTGCTGATTAATGCGCTGGGGTGCATCGGT 240

Qy 241 TACCGTTTATGTTTGGCTGCTGATGCGCGAAGAGCCATGACAAACAGAGTGAAGCG 300  
Db 241 TACCGTTTATGTTTGGCTGCTGATGCGCGAAGAGCCATGACAAACAGAGTGAAGCG 300

Qy 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTCCCGGATGACACGCTTTTGGGCG- GTTC 359  
Db 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTCCCGGATGACACGCTTTTGGGCG- GTTC 360

Qy 360 GCGGTGGGCACTGACACGCGCAACCGGTTGCGGTGCGAGTGAAGTGAAGTGAAGCG 419  
Db 360 GCGGTGGGCACTGACACGCGCAACCGGTTGCGGTGCGAGTGAAGTGAAGTGAAGCG 420

Qy 420 GGTTCGGTAAAGCCCATGTTGATGCGCAATGCTGCTGACCTGTTGAACCCGATGCGTA 479  
Db 420 GGTTCGGTAAAGCCCATGTTGATGCGCAATGCTGCTGACCTGTTGAACCCGATGCGTA 480

Qy 480 TTTTGGACGCGTTTGTGTTTATCGCGCGGCTCGCGCGCAATACGCGGACACCGGACG 539  
Db 480 TTTTGGACGCGTTTGTGTTTATCGCGCGGCTCGCGCGCAATACGCGGACACCGGACG 540

Qy 540 GATTTTCGCGCGTGGCGGTTTCGCGCAAGCTGATCTGCTGCTGCTGGTGGGTTTCGG 599  
Db 540 GATTTTCGCGCGTGGCGGTTTCGCGCAAGCTGATCTGCTGCTGCTGGTGGGTTTCGG 600

Qy 600 CGCAGCAGCATTTGTCAAGCCGCTGTCAGCGCCCAAGGTTGGCGCTGCGATCAACGTCGT 659  
Db 600 CGCAGCAGCATTTGTCAAGCCGCTGTCAGCGCCCAAGGTTGGCGCTGCGATCAACGTCGT 660

Qy 660 CGTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 711  
Db 660 CGTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 712

RESULT 9

AAH45375  
ID AAH45375 standard; DNA; 1568 BP.

XX AAH45375;

XX 11-SEP-2001 (first entry)





Db 13 TACTTTCAAGCTCTTGCTACCTGGCGGCTATGATCTACGCTCGCTCCACAAAATGCT 72  
 Qy 79 CTGGTGATTAACAAGGATTAAGCGGAAGGACTCATTCGGCTTCTTCGTGTGTTTA 138  
 Db 73 TTGTGTGATGATCAGGCGATACGCTGCTAGTACCAATATGATTTGCTGTGCT 132  
 Qy 139 ATTTCCTGACGCTTTTGTTCATCCCGGACCTTGGGGTGTGATCTTTTGTCCAAATGCC 198  
 Db 133 ATCAGCGATTGGTCTGATTTGCCCGGATTTTGGTGGCAGGGTATTGATGAG 192  
 Qy 199 GCGCCGATGCTGCTGATATATGCGCTGGGGTGGCATCGCTTACCTGTTATGTTGCC 258  
 Db 193 TCGCGGTGTTGCTGGCGCTGCTACCTGGGGCGGGTAGCTTCTTGTCTGTGTATGCT 252  
 Qy 259 GTCATGGCAGCAAGACGCCATGACAAACAGGTGGA 296  
 Db 253 TTTGGCGCTTTTAAACAGCAATGAGCAATATTTGA 290

RESULT 11  
 AAT33536/C  
 ID AAT33536 standard; DNA; 15239 BP.  
 XX AC AAT33536;  
 XX DT 16-OCT-2003 (revised)  
 XX DT 15-FEB-1998 (first entry)  
 XX DE BCG deletion region 2 and flanking sequences.

XX KW BCG delta 2; virulence; avirulence; attenuation; gene deletion;  
 XX KW mycobacteria; vaccine; infection; marker; ss.  
 XX OS Mycobacterium bovis; strain BCG.  
 XX FH Key Location/Qualifiers  
 FT misc\_feature 3382..14071  
 FT /\*tag= a  
 FT /note= "BCG delta 1 deletion region"  
 XX WO9625519-A1.  
 XX 22-AUG-1996.  
 XX 15-FEB-1996; 96WO-US001938.  
 XX 17-FEB-1995; 95US-00390878.  
 XX (PATH-) PATHOGENESIS CORP.  
 XX Stover CK, Mahairas GG;  
 XX WPI; 1996-393419/39.  
 XX Detecting markers for avirulence in Mycobacterium - used in production of  
 PT vaccines against bacterial infection, and to detect bacterial infection.  
 XX Example 1; Fig 2; 66pp; English.  
 XX This DNA sequence comprises Mycobacterium bovis BCG deletion sequence  
 CC BCGdelta2. A specific genetic deletion of this region results in an  
 CC avirulence phenotype of the mycobacterium. 2 Other deletion regions (see  
 CC AAT33535 and AAT33537) have also been detected. Identification involved  
 CC screening a BCG cosmid library with a radiolabeled probe obtained  
 CC following DNA subtraction between virulent Mycobacterium tuberculosis  
 CC H37Rv and avirulent BCG. The deletions provide useful markers for the  
 CC identification of an avirulent, or a virulent, mycobacterial phenotype.  
 CC Determination of avirulence requires the detection of the presence or  
 CC absence of the deletion; the deletions are detected either by detecting  
 CC the presence or absence of deletion junctions (see AAT33538-46), or by  
 CC detecting the presence or absence of the sequences contained within the  
 CC deletion. Deletion polypeptides are used as components of immunological  
 CC assays and in vaccines. (Updated on 16-OCT-2003 to standardise OS field)

XX SQ Sequence 15239 BP; 2878 A; 4791 C; 4729 G; 2841 T; 0 U; 0 Other;  
 Query Match 8.0%; Score 57; DB 2; Length 15239;  
 Best Local Similarity 53.1%; Pred. No. 4e-06;  
 Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;  
 Qy 431 AGCCCATGTTGATGCAATCGTCTGACCTGTTGAACCGCAATCGTATTGTGACGGCT 490  
 Db 4792 AGTCTCTGTGATCTGTGGCGCATTCAGTTCTTCAACCAACGCTTACCTCGACACCG 4733  
 Qy 491 TTGTGTTTATCGCGGGCTCGCGCGCAATACCGCGCACACCGGACGTTGATTTTCGCG 550  
 Db 4732 TCGTGTGTTAGTGGCGCTGGCCACGACGACACGCA--CCAGCGCTGGCTGTTCGGCC 4676  
 Qy 551 CTGGCGGCTTGGCGGCAAGCTGATCTGTTCCCGCTGGTGGGTTTCGGCGCAGGACAT 610  
 Db 4675 TCGGCGCGCTCACAGCCAGTGGGTATGGTTCCGCCACCTTCGGGTTTCGAGCGCGGCT 4616  
 Qy 611 TGTCAAGCCCGCTGTCCAGGCCCAAGGTGTGGCGCTGGATCAACGCTCGTGGCAGTTG 670  
 Db 4615 TGGCGGGCTGTTCACCAACCGCGCTGTGGAGAAATCCTCGACGGCTGATCGCGGTCA 4556  
 Qy 671 TGATGACCGCATTTGCCCATCAAACTGATGTTGA 703  
 Db 4555 TGATGTTGGCGTGGGAATCTCGTGCACGTGA 4523

RESULT 12  
 AAI99682\_22  
 Continuation (23 of 45) of AAI99682 from base 2200001 (Mycobacterium tuberculosis strai  
 WP Sequence split into 45 fragments LOCUS AAI99682 Accession Aai99682  
 WP Fragment Name Begin End  
 WP AAI99682\_00 1 110000  
 WP AAI99682\_01 100001 210000  
 WP AAI99682\_02 200001 310000  
 WP AAI99682\_03 300001 410000  
 WP AAI99682\_04 400001 510000  
 WP AAI99682\_05 500001 610000  
 WP AAI99682\_06 600001 710000  
 WP AAI99682\_07 700001 810000  
 WP AAI99682\_08 800001 910000  
 WP AAI99682\_09 900001 1010000  
 WP AAI99682\_10 1000001 1110000  
 WP AAI99682\_11 1100001 1210000  
 WP AAI99682\_12 1200001 1310000  
 WP AAI99682\_13 1300001 1410000  
 WP AAI99682\_14 1400001 1510000  
 WP AAI99682\_15 1500001 1610000  
 WP AAI99682\_16 1600001 1710000  
 WP AAI99682\_17 1700001 1810000  
 WP AAI99682\_18 1800001 1910000  
 WP AAI99682\_19 1900001 2010000  
 WP AAI99682\_20 2000001 2110000  
 WP AAI99682\_21 2100001 2210000  
 WP AAI99682\_22 2200001 2310000  
 WP AAI99682\_23 2300001 2410000  
 WP AAI99682\_24 2400001 2510000  
 WP AAI99682\_25 2500001 2610000  
 WP AAI99682\_26 2600001 2710000  
 WP AAI99682\_27 2700001 2810000  
 WP AAI99682\_28 2800001 2910000  
 WP AAI99682\_29 2900001 3010000  
 WP AAI99682\_30 3000001 3110000  
 WP AAI99682\_31 3100001 3210000  
 WP AAI99682\_32 3200001 3310000  
 WP AAI99682\_33 3300001 3410000  
 WP AAI99682\_34 3400001 3510000  
 WP AAI99682\_35 3500001 3610000  
 WP AAI99682\_36 3600001 3710000  
 WP AAI99682\_37 3700001 3810000  
 WP AAI99682\_38 3800001 3910000  
 WP AAI99682\_39 3900001 4010000



Fragment Name	Begin	End
AAI99682_00	1	110000
WP		
AAI99682_01	100001	210000
WP		
AAI99682_02	200001	310000
WP		
AAI99682_03	300001	410000
WP		
AAI99682_04	400001	510000
WP		
AAI99682_05	500001	610000
WP		
AAI99682_06	600001	710000
WP		
AAI99682_07	700001	810000
WP		
AAI99682_08	800001	910000
WP		
AAI99682_09	900001	1010000
WP		
AAI99682_10	1000001	1110000
WP		
AAI99682_11	1100001	1210000
WP		
AAI99682_12	1200001	1310000
WP		
AAI99682_13	1300001	1410000
WP		
AAI99682_14	1400001	1510000
WP		
AAI99682_15	1500001	1610000
WP		
AAI99682_16	1600001	1710000
WP		
AAI99682_17	1700001	1810000
WP		
AAI99682_18	1800001	1910000
WP		
AAI99682_19	1900001	2010000
WP		
AAI99682_20	2000001	2110000
WP		
AAI99682_21	2100001	2210000
WP		
AAI99682_22	2200001	2310000
WP		
AAI99682_23	2300001	2410000
WP		
AAI99682_24	2400001	2510000
WP		
AAI99682_25	2500001	2610000
WP		
AAI99682_26	2600001	2710000
WP		
AAI99682_27	2700001	2810000
WP		
AAI99682_28	2800001	2910000
WP		
AAI99682_29	2900001	3010000
WP		
AAI99682_30	3000001	3110000
WP		
AAI99682_31	3100001	3210000
WP		

Job time : 293.925 sec

Job time : 293.925 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:47:35 ; Search time 53.5608 Seconds  
(without alignments)  
7366.769 Million cell updates/sec

Title: US-09-105-117K-1\_COPY\_1016\_1726

Perfect score: 711  
Sequence: 1 atgggtgatcgagaaattttt.....aactgatgtgatgggttag 711

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	61.8	8.7	699	4	US-09-252-991A-7908	Sequence 7908, Ap
2	61.8	8.7	834	4	US-09-252-991A-7643	Sequence 7643, Ap
3	61.8	8.7	894	4	US-09-252-991A-7841	Sequence 7841, Ap
4	57	8.0	597	4	US-09-894-844-11	Sequence 11, Appl
5	57	8.0	15239	1	US-08-390-878-17	Sequence 17, Appl
6	57	8.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl
7	57	8.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl
8	56.6	8.0	642	4	US-09-489-039A-4674	Sequence 4674, Ap
9	53.8	7.6	1176	4	US-09-252-991A-7571	Sequence 7571, Ap
10	43	6.0	645	4	US-09-543-681A-800	Sequence 800, App
11	39.4	5.5	505	4	US-09-621-976-15639	Sequence 800, App
12	37.6	5.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
13	37.6	5.3	4411529	3	US-09-103-840A-1	Sequence 1, Appl
14	35.4	5.0	636	4	US-09-328-352-1611	Sequence 1611, Ap
15	35	4.9	1002	4	US-09-252-991A-1099	Sequence 1099, Ap
16	35	4.9	1092	4	US-09-252-991A-946	Sequence 946, Ap
17	35	4.9	1521	4	US-09-252-991A-989	Sequence 989, App
18	34.8	4.9	723	4	US-09-252-991A-1130	Sequence 1130, Ap
19	34.8	4.9	1134	4	US-09-252-991A-1056	Sequence 1056, Ap
20	34.8	4.9	1233	4	US-09-252-991A-2121	Sequence 2121, Ap
21	34.8	4.9	1254	4	US-09-252-991A-904	Sequence 904, App
22	34.8	4.9	1395	4	US-09-252-991A-1954	Sequence 1954, App
23	34.6	4.9	417	4	US-09-252-991A-1351	Sequence 1351, Ap
24	34.6	4.9	1029	4	US-09-252-991A-1752	Sequence 1752, Ap
25	34.6	4.9	1098	4	US-09-252-991A-1304	Sequence 1304, Ap
26	34.6	4.9	1215	4	US-09-252-991A-1975	Sequence 1975, Ap
27	34.6	4.9	1230	4	US-09-252-991A-2136	Sequence 2136, Ap

28	34.6	4.9	1632	4	US-09-252-991A-1216	Sequence 1216, Ap
29	34.6	4.9	1746	4	US-09-252-991A-1163	Sequence 1163, Ap
30	34.6	4.9	1995	4	US-09-252-991A-14545	Sequence 14545, A
31	34.6	4.9	3396	4	US-09-252-991A-14676	Sequence 14676, A
32	34.6	4.9	3444	4	US-09-252-991A-15078	Sequence 15078, A
33	34.6	4.9	6611	1	US-08-402-282-2	Sequence 2, Appl
34	34.6	4.9	6611	1	US-08-508-004-2	Sequence 2, Appl
35	34.6	4.9	6611	1	US-08-402-066-2	Sequence 2, Appl
36	34.6	4.9	6611	1	US-08-402-068-2	Sequence 2, Appl
37	34.2	4.8	495	4	US-09-252-991A-15856	Sequence 15856, A
38	34.2	4.8	804	4	US-09-252-991A-15809	Sequence 15809, A
39	34.2	4.8	963	4	US-09-489-039A-5986	Sequence 5986, Ap
40	34.2	4.8	2073	4	US-09-252-991A-15779	Sequence 15779, A
41	33.8	4.8	1284	4	US-09-252-991A-13377	Sequence 13377, A
42	33.8	4.8	1377	4	US-09-252-991A-13756	Sequence 13756, A
43	33.8	4.8	2318	4	US-09-620-312D-733	Sequence 733, App
44	33.6	4.7	717	4	US-09-252-991A-13565	Sequence 13565, A
45	33.6	4.7	1452	4	US-09-252-991A-10581	Sequence 10581, A

ALIGNMENTS

RESULT 1

US-09-252-991A-7908  
; Sequence 7908, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7908  
; LENGTH: 699  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7908

Query Match	8.7%;	Score 61.8;	DB 4;	Length 699;
Best Local Similarity	54.6%;	Pred. No. 4.9e-09;		
Matches	148;	Conservative	0;	Mismatches 117;
				Indels 6;
				Gaps 1;
Qy	440	TGATGGCAATCGTGTGACCTGGTGTGAACCCGAATGCGTATTTGGACGGTTTGTTTA	499	
Db	434	TGGCGGCACTGGCGGTACCTGCTCAACCCCACTCTATCTCGATACCGTCTCTGA	493	
Qy	500	TGGCGGCGTGGCGGCAACACCGGACACCGGCGTGGATTTTCGGCGTGGCGGT	559	
Db	494	TGGTTCGTCGCGGCGCCACGAGG-----CGCGCGGCGGCTATGCCCTCGCGCGG	547	
Qy	560	TGGCGGCAACCTGATCTGTGTTCCCGCTGGTGGTTCGGCGGAGCAGCATTTGACGCC	619	
Db	548	CCAGCGCTCGCTGATGTTCTTCGCGCTCGCCCTCGCGCGGCGATGGCTGGCGCCCT	607	
Qy	620	CGCTGTTCAGCCCGCAAGGTGTGGCGCTGGATCAAGCTGTGGCGAGTTGTGATGACCG	679	
Db	608	GGCTGGCGCGCGCGCCACCTGGCGGCTCTGCTGACCTGATGGTGGCGGCGATGATGCTGG	667	
Qy	680	CATTGGCCATCAACTGATGTTGATGGTTA	710	
Db	668	GCATGGCGGCGCACTGCTGTTCCGGGATA	698	

RESULT 2

US-09-252-991A-7643/c  
; Sequence 7643, Application US/09252991A

	Query Match	8.7%	Score 61.8;	DB 4;	Length 894;
	Best Local Similarity	54.6%;	Pred. No. 5.5e-09;		
	Matches 148;	Conservative	0;	Mismatches 117;	Indels 6; Gaps 17;
QY	440	TGATGGCAATGCTCTGACCTCGTTTGAACCGAATGCGTATTTGGACGCGTTTGTCTTTA	499		
ph	339	TGGCGCATTTGGCGTACCTCTGTCGAACCCCGACGTCATCTTCGATACCGTACTCTCTGA	398		

Query Match	8.0%	Score 57	DB 4	Length 597
Best Local Similarity	53.1%	Pred. No. 1.3e-07		
Matches 145	Conservative 0	Mismatches 125	Indels 3	Gaps 17
QY	431	AGCCCATGTTATGCCAATCGTGTCTGACCTGTGGTGTGAACCCGAAATGCGTATTTGGACGGCT	490	
DB	326	AGGTCCTGTGACCTGTGGCGCATTCACGTTCTCTCAACCCACACGCTACCTCGACACCG	385	
QY	431	TTGTGTTTATCGCGGGCTGCGCGGCAATACGGGGACACGGACGCTGGATTTTTCGCG	550	
DB	386	TCGTGTGTCTAGCGCGCGCTGGCCCAACGAGCACAGCA---CCAGCGCTGGCTGTTCCGCC	442	
QY	551	CTGCGCGCTTTCGCGCGAAGCTGATCTCTGTTTCCCGCTGGTGGGTTTTCGCGCGCACGACAT	610	
DB	443	TCGCGCGCTCACACGCCAGTCGGGTATGTTTCCACCCTTCGGGTTTCGGAGCGCGGCGGCT	502	
QY	611	TGTCACGCGCGCTGCCACGCCCAAGGTGTGGCGCTGGATCAACGTCGTCGTGGCAGCTTG	670	
DB	503	TGCGCGGGCTGTTTACCAACACCCCGCTCTGTGGAAATCTCTGACGCGCTGATCCGGGTCA	562	
QY	671	TGATGACCGCATTCGGCCATCAAACTGATGTGA	703	
DB	563	TGATGTTGCGCTGGGAATCTCGCTGACCGTGA	595	

RESULT 5  
US-08-390-878-17/c  
; Sequence 17, Application US/08390878  
; Patent No. 5700583

## GENERAL INFORMATION:

APPLICANT: Scover, Charles K.  
APPLICANT: Mahairas, Gregory G.  
TITLE OF INVENTION: VIRULENCE-ATTENUATING GENETIC DELETIONS  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Townsend and Townsend Kourile and Crew  
STREET: One Market Plaza, Steuart Street Tower, 20th  
STREET: Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/390,878  
FILING DATE: 17-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunter, Tom  
REGISTRATION NUMBER: 38,498  
REFERENCE/DOCKET NUMBER: 15371A-17  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/543/9600  
TELEFAX: 415/543/5043  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15239 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-390-878-17

Query Match 8.0%; Score 57; DB 1; Length 15239;  
Best Local Similarity 53.1%; Pred. No. 6.6e-07;  
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 431 AGCCCATGTTGATGGCAATCGTCTGACCTGGTTGAACCCGAATCGGTATTGGACGGT 490  
DB 4792 AGTCTCTGTGACCTGTGGCGATTTCAGTTCTCAACCCACACGCTTACCTCGACACG 4733  
QY 491 TTGTGTTATCGCGCGGTCTGGCGGCAATACGGCGACACCGGACGGTGGATTTCGCG 550  
DB 4732 TCGTGTGTTAGCGCGCTGGCCAAACGACACAGCGA--CCAGCGCTGGCTTTCGGCC 4676  
QY 551 CTGGCGGTTCGGCGAGCGCTGATCTGTTCCCGCTGGTGGTTTGGCGGACGACGAT 610  
DB 4675 TCGCGCGGTCTACAGCCAGTGGCGGTATGTTTCGCCACCTCGGGTTCGAGCGCGCGGT 4616  
QY 611 TGTACGCGCGCTGTCCAGCCCCCAAGTGTGGCGCTGGATCAACGTCGTGGCGAGTTG 670  
DB 4615 TGGCGGGGTGTTTCAACCAACCCCGGCTGTGGAGAACTCTCGACGCGCTGATCGCGTCA 4556  
QY 671 TGATACCGCATTTGGCCATCAAACTGATGTTGA 703  
DB 4555 TGATGTTGCGTGGGAATCTCGCTGACCGTGA 4523

## RESULT 6

US-09-103-840A-2  
Sequence 2, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

## TITLE OF INVENTION: TUBERCULOSIS

FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 2  
LENGTH: 4403765  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
FEATURE:  
OTHER INFORMATION: CDC 1551  
OTHER INFORMATION: "n" bases at various positions throughout the sequence  
OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 8.0%; Score 57; DB 3; Length 4403765;

Best Local Similarity 53.1%; Pred. No. 1.1e-05;  
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 431 AGCCCATGTTGATGGCAATCGTCTGACCTGGTTGAACCCGAATCGGTATTGGACGGT 490  
DB 2227633 AGTCTCTGTGACCTGTGGCGATTTCAGTTCTCAACCCACACGCTTACCTCGACACG 2227692  
QY 491 TTGTGTTATCGCGCGGTCTGGCGGCAATACGGCGACACCGGACGGTGGATTTCGCG 550  
DB 2227693 TCGTGTGTTAGCGCGCTGGCCAAACGACACAGCGA--CCAGCGCTGGCTTTCGGCC 2227749  
QY 551 CTGGCGGTTCGGCGAGCGCTGATCTGTTCCCGCTGGTGGTTTGGCGGACGACGAT 610  
DB 2227750 TCGGCGCGGTCTACAGCCAGTGGCGGTATGTTTCGCCACCTTCGGGTTTCGAGCGCGGT 2227809  
QY 611 TGTACGCGCGCTGTCCAGCCCCCAAGTGTGGCGCTGGATCAACGTCGTGGCGAGTTG 670  
DB 2227810 TGGCGGGGTGTTTCAACCAACCCCGGCTGTGGAGAACTCTCGACGCGCTGATCGCGTCA 2227869  
QY 671 TGATACCGCATTTGGCCATCAAACTGATGTTGA 703  
DB 2227870 TGATGTTGCGTGGGAATCTCGCTGACCGTGA 2227902

## RESULT 7

US-09-103-840A-1  
Sequence 1, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 8.0%; Score 57; DB 3; Length 4411529;

Best Local Similarity 53.1%; Pred. No. 1.1e-05;  
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 431 AGCCCATGTTGATGGCAATCGTCTGACCTGGTTGAACCCGAATCGGTATTGGACGGT 490  
DB 2230334 AGTCTCTGTGACCTGTGGCGATTTCAGTTCTCAACCCACACGCTTACCTCGACACG 2230393  
QY 491 TTGTGTTATCGCGCGGTCTGGCGGCAATACGGCGACACCGGACGGTGGATTTCGCG 550

Db 2230394 TCCTGTTGCTAGGCGCGTGTCCCAACAGCAGCAGCA ---CAGCGCTGCTGTTGGCC 2230450  
 QY 551 CTGCGCGGTTTCGCGCAAGCTGATCTGTTCCCGCTGGTGGGTTTCGCGCAGCAGCAT 610  
 Db 2230451 TCAGCGGCTGACAGCAGTGGGTATGTTCCGACCTCGGGTTGCGAGCGCGCGGT 2230510  
 QY 611 TGTCAACCGCGCTGTCAGCGCCCAAGGTGTGGCGCTGGATCAACGTCGCTGCGCAGTTG 670  
 Db 2230511 TCGCGCGGCTGTTTCCCAACCGCGCTCGTGGAGATCTCTCGACGCGCTGATCCCGGTCA 2230570  
 QY 671 TCAGCGCGCATGGCCATCAAACTGATGTGA 703  
 Db 2230571 TGATGGTTCGCTGGGAATCTCGCTGACCGTGA 2230603

RESULT 8

US-09-489-039A-4674  
 ; Sequence 4674, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 2709.2004001  
 ; CURRENT APPLICATION NUMBER: US/09/489,039A  
 ; PRIOR FILING DATE: 2000-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/117,747  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 4674  
 ; LENGTH: 642  
 ; TYPE: DNA  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-4674

Query Match 8.0%; Score 56.6; DB 4; Length 642;  
 Best Local Similarity 50.2%; Pred. No. 1.8e-07;  
 Matches 140; Conservative 0; Mismatches 139; Indels 0; Gaps 0;  
 QY 19 TTCAATCAGCTCTGCTTTTGGGGCCAGCTTTTACTGTCCATCGGACCGCAGAAATGA 78  
 Db 19 TACTTTCAGGGCTTCCCTTGGAGGCGCTATGATATGCGCTTGGCCCAACAAATGCC 78  
 QY 79 CTGCGATTAACAGGAATTAAGCGAAGACTCATTTGGGTTCTTCTCGTGTTGA 138  
 Db 79 TTGTGATGAATCAGGGGATCCGTCGTGATCATCATCTGATGATGCGCTGCTATGCCG 138  
 QY 139 ATTCTGACGCTTTTGTTCATCGCGCACCTTGGCGTTGATCTTTTGTCCAAATGCC 198  
 Db 139 GTCAGGACTTGTGTTAATCTGCGTGGCATTTTGGCGGTAGCGCACTGCTGATGCA 198  
 QY 199 GCGCGCATGCTGCTGATATTAAGCTGGGTTGGCATCGCTTACCTGTTATGTTTCC 258  
 Db 199 TCCCGTGGCTGCTGGCGCTAGTACCTTGGCGCGGTGGCGCTTTTGTCTGTGTATGCC 258  
 QY 259 GTACGCGAGCGAAGACCGCATGACAAACAGTCTGA 297  
 Db 259 TTGCGCGAATTGAAACGCGGTTTCAGTCAGAGTCTTGA 297

RESULT 9

US-09-252-991A-7571/c  
 ; Sequence 7571, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 7571  
 ; LENGTH: 1176  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-7571

Query Match 7.6%; Score 53.8; DB 4; Length 1176;  
 Best Local Similarity 54.3%; Pred. No. 1.7e-06;  
 Matches 134; Conservative 0; Mismatches 107; Indels 6; Gaps 1;  
 QY 464 TGAACCCGAATGCTATTTTGACGCGGTTTGTGTTTATCGCGCGCTCGCGCGCAATACG 523  
 Db 1176 TCARCCCCACGTCATCTCATACCGTACCTTCTGATCGGTTGCTGCGCGCCAGCAGG 1117  
 QY 524 GCGACACCGGACGCTGGAATTTTCGCGCTGCGCGCTGCGCGCAAGCTGATCTGTTCC 583  
 Db 1116 -----CCGCGCGCGCGCTATGCTCGCGCGCGCGCGCGCTCGCTGATGTTCT 1063  
 QY 584 CGCTGCGGTTTTCGCGCGCAGCAGTATGACGCGCGCTGTCAGCGCCCAAGTGTGCC 643  
 Db 1062 TCGCCCTCGCCCTCGCGCGCGATGCTGCGCCCTTGGCTGCGCGCCCGGCCACCTGCC 1003  
 QY 644 GCTGATCAACGCTGCTGCGCGCTGCTGATGACCGCATTTGGCCATCAAACTGATTTGA 703  
 Db 1002 GCTGCTGACCTGATGCTGCGCGCATGATGCTGCGCATGCTGCGCATGCTGCTTCC 943  
 QY 704 TGGGTTA 710  
 Db 942 GCGGATA 936

RESULT 10

US-09-543-681A-800  
 ; Sequence 800, Application US/09543681A  
 ; Patent No. 6605709  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GARY BRETON  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRAB.  
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 2709.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/543,681A  
 ; PRIOR FILING DATE: 2000-04-05  
 ; PRIOR APPLICATION NUMBER: US 60/128,706  
 ; PRIOR FILING DATE: 1999-04-09  
 ; NUMBER OF SEQ ID NOS: 8344  
 ; SEQ ID NO 800  
 ; LENGTH: 645  
 ; TYPE: DNA  
 ; ORGANISM: Proteus mirabilis  
 US-09-543-681A-800

Query Match 6.0%; Score 43; DB 4; Length 645;  
 Best Local Similarity 46.7%; Pred. No. 0.0025;  
 Matches 136; Conservative 0; Mismatches 155; Indels 0; Gaps 0;  
 QY 15 AATCTTCATACAGCTCTGCTTTTGGGGCCAGTCTTTTACTGTCCATCGCAGCGCAGAA 74  
 Db 30 AACATTTTTCAGGGGTTTATTAGTGCGAGCATGATTTGCCCATTTGGTGGCAGAA 89  
 QY 75 TGATCTGTTGATTAACAGGAATTAAGCGGAGGACCTATTCGGGTTCTTCTCGTGG 134  
 Db 90 TGCCTTTGTTTGCACAGGTAGTAAACAAATACCATTTAATAGTGCATTTTATG 149  
 QY 135 TTTAATTTCTGAGCTCTTTTGTTCATCGCGCACCTTGGCGTGTGATCTTTTGTCCA 194  
 Db 150 CGCTTTAAGTGATGTGATATTAATTAAGCGGGGAGTCTTTTGGTGAAGTGATTTAG 209  
 QY 195 TCGCGCGCGATCGTCTCGATATTTATGCTGGGGTGGCATGCTTACCTGTATGTT 254



Db 210 CCAATCAGAAATATCTTTTATTACTGATCACTTGGGCGGTGGCTTTTATTGTCGTA 269  
 Qy 255 TCCCGTCATGCGAGCAAGACGCCATGACAAACCAAGGTGGAAGCCACCA 305  
 Db 270 TGGATGCAATGCGTTTAAACCGCTTTTACAAAGAGATTGAATTATCACA 320

RESULT 11

US-09-621-976-15639  
 ; Sequence 15639, Application US/09621976  
 ; Patent No. 6639063  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Jober, S.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET.054PR2  
 ; CURRENT APPLICATION NUMBER: US/09/621,976  
 ; CURRENT FILING DATE: 2000-07-21  
 ; NUMBER OF SEQ ID NOS: 19335  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 15639  
 ; LENGTH: 505  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-621-976-15639

Query Match 5.5%; Score 39.4; DB 4; Length 505;  
 Best Local Similarity 11.4%; Pred. No. 0.027;  
 Matches 28; Conservative 118; Mismatches 99; Indels 0; Gaps 0;  
 Qy 379 CGCAACCGGTGGCGGTGAGGTGAGCGTCGATACGACGAGCGGTGGGTAAGCCCAAG 438  
 Db 5 MSRYRTSSKRYGKYSMYSSRWKMYANGRYGTSGRCGRGSMCKWGYRS 64  
 Qy 439 TTGATGCAATCTGCTGACCTGCTGTAACCGAATGATTTTGACGCGTTTGCTTT 498  
 Db 65 VAGYKWSKSKMYKSGMTGSKTKRTYTSKRTYTKYRGMWSKWRMWRKMY 124  
 Qy 499 ATCGCGCGGTGCGCGCGCAATACGCGACACCGAGCGGTGGATTTTCGCGCTGGCGG 558  
 Db 125 YRMKCYSCASYSYRRCRYTGMTRGMYGCKRMCKSSRYMTRYKMYGACYGS 184  
 Qy 559 TTCGCGCAAGCTGATCTGTTCCCGCTGGTGGGTTTCGGCGCAGCAGCATTTGTCACG 618  
 Db 185 KMSCKGSRKYGKWKYGYKTYMCTSKYKSKMYKSKMYCYTYMCMYTYWMCCTY 244  
 Qy 619 CCGCT 623  
 Db 245 KSYI 249

RESULT 12

US-09-103-840A-2/c  
 ; Sequence 2, Application US/09103840A  
 ; Patent No. 6294328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEISCHMAN, Robert D.  
 ; APPLICANT: WHITE, Owen R.  
 ; APPLICANT: FRASER, Claire M.  
 ; APPLICANT: VENTER, John C.  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
 ; FILE REFERENCE: TUBERCULOSIS  
 ; CURRENT APPLICATION NUMBER: US/09/103,840A  
 ; CURRENT FILING DATE: 1998-06-24  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 4403765  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; FEATURE:

OTHER INFORMATION: CDC 1551  
 ; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
 ; OTHER INFORMATION: represent a, t, c or g  
 US-09-103-840A-2  
 Query Match 5.3%; Score 37.6; DB 3; Length 4403765;  
 Best Local Similarity 47.5%; Pred. No. 8.7;  
 Matches 112; Conservative 0; Mismatches 124; Indels 0; Gaps 0;  
 Qy 430 AAGCCCATGTTGATGCAATGCTGCTGACCTGTTGAACCCGCAATGCTATTGTCAGCG 489  
 Db 720443 AAGGACCTGCGCTGCGCTGCGCTGCGCGCGCGCGGAGCTGCCCGGCCCGCTCGGT 720384  
 Qy 490 TTTGTGTTTATCGCGCGGTGCGCGCGCAATACGCGCAGCACCGGACGGTGGATTTCGCC 549  
 Db 720383 GGGCTGCTGTTTACCTGTCAACGGCGCGGACGACGAATGTTTCGGGGTCAACGACGAC 720324  
 Qy 550 GCTGGCGGTTCGCGCAAGCTGATCTGTTCCCGCTGGTGGGTTTCGGCGCAGCAGCA 609  
 Db 720323 GCGTCGAGATCGAGGACCTGCTGGCGGGAATTCGCTGGCGCGGTTTCTTCGCGCGCGG 720264  
 Qy 610 TTGTCAGCGCGCTGCTCAGCCCAAGGCTGTCGCGCTGATCAACGTCGTCGTGC 665  
 Db 720263 GAGATCGCGCGGTGCGCGGCAACAGCGCTTGACGCGGTTTACCGGTCGATGC 720208

RESULT 13

US-09-103-840A-1/c  
 ; Sequence 1, Application US/09103840A  
 ; Patent No. 6294328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEISCHMAN, Robert D.  
 ; APPLICANT: WHITE, Owen R.  
 ; APPLICANT: FRASER, Claire M.  
 ; APPLICANT: VENTER, John C.  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
 ; FILE REFERENCE: TUBERCULOSIS  
 ; CURRENT APPLICATION NUMBER: US/09/103,840A  
 ; CURRENT FILING DATE: 1998-06-24  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 4411529  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; OTHER INFORMATION: H37Rv  
 US-09-103-840A-1

Query Match 5.3%; Score 37.6; DB 3; Length 4411529;  
 Best Local Similarity 47.5%; Pred. No. 8.7;  
 Matches 112; Conservative 0; Mismatches 124; Indels 0; Gaps 0;  
 Qy 430 AAGCCCATGTTGATGCAATGCTGCTGACCTGTTGAACCCGCAATGCTATTGTCAGCG 489  
 Db 719010 AAGGACCTGCGCTGCGCTGCGCTGCGCGCGCGGAGCTGCCCGGCCCGCTCGGT 718951  
 Qy 490 TTTGTGTTTATCGCGCGGTGCGCGCGCAATACGCGCAGCACCGGACGGTGGATTTCGCC 549  
 Db 718950 GGGCTGCTGTTTACCTGTCAACGGCGCGGACGACGAATGTTTCGGGGTCAACGACGAC 718891  
 Qy 550 GCTGGCGGTTCGCGCAAGCTGATCTGTTCCCGCTGGTGGGTTTCGGCGCAGCAGCA 609  
 Db 718890 GCGTCGAGATCGAGGACCTGCTGGCGGGAATTCGCTGGCGCGGTTTCTTCGCGCGCGG 718831  
 Qy 610 TTGTCAGCGCGCTGCTCAGCCCAAGGCTGTCGCGCTGATCAACGTCGTCGTGC 665  
 Db 718830 GAGATCGCGCGGTGCGCGGCAACAGCGCTTGACGCGGTTTACCGGTCGATGC 718775

RESULT 14

US-09-328-352-1611  
 ; Sequence 1611, Application US/09328352

Thu Mar 18 12:31:01 2004

Patent No. 6562958  
 GENERAL INFORMATION:  
 APPLICANT: Gary L. Breton et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: GTC99-03PA  
 CURRENT APPLICATION NUMBER: US/09/328,352  
 CURRENT FILING DATE: 1999-06-04  
 NUMBER OF SEQ ID NOS: 8252  
 SEQ ID NO 1611  
 LENGTH: 636  
 TYPE: DNA  
 ORGANISM: Acinetobacter baumannii  
 US-09-328-352-1611

Query Match 5.0%; Score 35.4; DB 4; Length 636;  
 Best Local Similarity 48.7%; Pred. No. 0.5;  
 Matches 96; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
 QY 17 TCTTCATTACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTCCATCGGACCGCAGATG 76  
 DB 50 TGTCTTTTAAAGGTTTAGGTATGTGTAGCGGCTTAATTGTAGCAATTGGTGGCCAAAATG 109  
 QY 77 TACTGGTGATTAACAGCAATTAAGCGCGAGGACTCATTTGGGTTCTTCTCGTGTGTT 136  
 DB 110 CTTTGTGTTCTAAAGCAAGGGCTTAAACAGCAATATGATTTTGGTTATGTTTAACTGTG 169  
 QY 137 TAATTTCTGAGCTCTTTTGTTCATCGCGGACCTTGGCGTTGATCTTTTGTCCATG 196  
 DB 170 CCTCTCGGATCAATTTAATCGCTGTGTGTTAGGTTTTCGCGAAATTATGACAG 229  
 QY 197 CGCGCGCATCGTGCTC 213  
 DB 230 CATCACCATCTTAATC 246

RESULT 15  
 US-09-252-991A-1099  
 Sequence 1099, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 1099  
 LENGTH: 1002  
 TYPE: DNA  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-1099

Query Match 4.9%; Score 35; DB 4; Length 1002;  
 Best Local Similarity 48.3%; Pred. No. 0.83;  
 Matches 131; Conservative 0; Mismatches 135; Indels 5; Gaps 1;  
 QY 420 GGTTCGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGTTGATCCCGCAATCGTA 479  
 DB 613 GGCCTTGTATCGACCGAGCAGAGGTTCATCATCCCGAGGTCCATCGAGTCAAGAGAA 672  
 QY 480 TTGAGACGCTTTGTGTTTATCGCGCGTTCGCGCAATACGCGACACCGACCGTG 539  
 DB 673 CGCCAGCGCGATGATCAAGAGACAGTCTCGGTATACGACTCAGCGGACGCGTTCGAG 732  
 QY 540 GATTTTCGCGGCGGCTTCGCGGCAAGCCTGATCTGTTCCTCCGCTGGTTCGGTTCCG 599  
 DB 733 GCGTTCGCGCGGATGCGGCGGTTCGCGCATGACCTGTGTTCTCTGCGCGCGGCGGT 792

QY 600 CGCAGCAGCATTTGTCAAGCCCGCTGTCCAGCCCAAGGTGTGGCGCTGGATCAACGTCGT 659  
 DB 793 CGAACGAGTCTGGATCGCC-----GCCCGCGCGTGTGCTTGGCGCGCAACCGCGGT 847  
 QY 660 CGTGGCAGTTGTGATGACCGCGCATTTGGCCATC 690  
 DB 848 CGAGCCGATTGGCAGCCCGCGCGCTGGCCGTC 878

Search completed: March 16, 2004, 03:00:56  
 Job time : 78.5608 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2004, 00:51:03 ; Search time 254.891 Seconds  
(without alignments)  
10269.549 Million cell updates/sec

Title: US-09-105-117K-1\_COPY\_1016\_1726

Perfect score: 711  
Sequence: 1 atgttgatcaggaatctt.....aactgatgttgaggtag 711

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 243257 seqs, 184079884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	711	100.0	711	14	US-10-166-142-7
2	711	100.0	822	10	US-09-746-660A-51
3	711	100.0	2374	14	US-10-196-232-24
4	711	100.0	3309400	9	US-09-738-626-1
5	708	99.6	708	9	US-09-738-626-3455
6	696.8	98.0	712	14	US-10-166-142-9
7	62.8	8.8	609	14	US-10-156-761-1109
8	62.8	8.8	9025608	14	US-10-156-761-1
9	57	8.0	597	9	US-09-894-844-11
10	57	8.0	597	15	US-10-388-902-11
11	56.2	7.9	86114	14	US-10-080-170-648
12	42.2	5.9	3000	14	US-10-156-761-5949
13	42.2	5.9	9025608	14	US-10-156-761-1
14	40.4	5.7	1392	15	US-10-369-493-31368
15	40.4	5.7	1443	15	US-10-369-493-28610

c 16	39.4	5.5	594	15	US-10-369-493-33277	Sequence 33277, A
c 17	39.2	5.5	1257	14	US-10-156-761-5999	Sequence 5999, Ap
c 18	39	5.5	993	10	US-09-746-660A-47	Sequence 47, Appl
c 19	39	5.5	2823	12	US-10-282-122A-26243	Sequence 26243, A
c 20	39	5.5	2826	12	US-10-282-122A-28428	Sequence 28428, A
c 21	38.6	5.4	1245	12	US-10-425-114-960	Sequence 960, Ap
c 22	38	5.3	1018	12	US-10-425-114-960	Sequence 14346, A
c 23	38	5.3	1458	14	US-10-156-761-1306	Sequence 1306, Ap
c 24	38	5.3	2218	12	US-10-425-114-28221	Sequence 28221, A
c 25	38	5.3	2457	12	US-10-425-114-30614	Sequence 30614, A
c 26	37.8	5.3	536	16	US-10-338-110-119	Sequence 119, App
c 27	37.8	5.3	2038	14	US-10-094-113-11	Sequence 11, Appl
c 28	37.6	5.3	536	16	US-10-338-110-119	Sequence 119, App
c 29	37.6	5.3	2079	14	US-10-156-761-1755	Sequence 1755, Ap
c 30	37.6	5.3	3073	12	US-10-424-599-33204	Sequence 33204, A
c 31	37.6	5.3	3090	12	US-10-425-114-11304	Sequence 11304, A
c 32	37.4	5.3	624	14	US-10-156-761-2439	Sequence 2439, Ap
c 33	37.4	5.3	1218	15	US-10-369-493-44172	Sequence 44172, A
c 34	37.4	5.3	1350	12	US-10-282-122A-15180	Sequence 15180, A
c 35	37.4	5.3	1631	15	US-10-369-493-33258	Sequence 33258, A
c 36	37.2	5.2	424	9	US-09-960-352-12392	Sequence 12392, A
c 37	37.2	5.2	873	12	US-10-282-122A-25710	Sequence 25710, A
c 38	37	5.2	945	12	US-10-282-122A-14462	Sequence 14462, A
c 39	37	5.2	1410	14	US-10-156-761-3506	Sequence 3506, Ap
c 40	37	5.2	1464	12	US-10-282-122A-31748	Sequence 31748, A
c 41	36.8	5.2	546	14	US-10-156-761-4089	Sequence 4089, Ap
c 42	36.8	5.2	678	12	US-10-282-122A-26365	Sequence 26365, A
c 43	36.8	5.2	681	12	US-10-282-122A-28183	Sequence 28183, A
c 44	36.8	5.2	954	12	US-10-282-122A-26045	Sequence 26045, A
c 45	36.6	5.1	681	15	US-10-260-238-3605	Sequence 3605, Ap

ALIGNMENTS

RESULT 1

US-10-166-142-7  
; Sequence 7, Application US/10166142  
; Publication No. US20030124687A1  
; GENERAL INFORMATION:  
; APPLICANT: GUNJI, YOSHIYA  
; APPLICANT: YASUEDA, HISASHI  
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL  
; FILE REFERENCE: 223789US  
; CURRENT APPLICATION NUMBER: US/10/166,142  
; CURRENT FILING DATE: 2002-06-11  
; PRIOR APPLICATION NUMBER: JP 2001-1777075  
; PRIOR FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 7  
; LENGTH: 711  
; TYPE: DNA  
; ORGANISM: Brevibacterium lactofermentum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(711)  
; OTHER INFORMATION:  
US-10-166-142-7

Query Match	100.0%	Score	711	DB	14	Length	711
Best Local Similarity	100.0%	Pred. No.	3.7e-225				
Matches	711	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	ATGTTGATCATCGAAATCTTCATTACAGTCTGCTTTTGGGGGCCAGCTCTTTTACTGTCC	60				
Db	1	ATGTTGATCATCGAAATCTTCATTACAGTCTGCTTTTGGGGGCCAGCTCTTTTACTGTCC	60				
Qy	61	ATCGACCCGACAGATGTACTGGTATTAACAGAAATTAAGCGGAGGACTCATTTGCG	120				
Db	61	ATCGACCCGACAGATGTACTGGTATTAACAGAAATTAAGCGGAGGACTCATTTGCG	120				

Qy	121	GTTCCTCTCGTGTGTTAAATCTTGACGCTCTTTTGTTCATPCGCGCGCACCTTGGCGTT	180
Db	121	GTTCCTCTCGTGTGTTTAAATCTTGACGCTCTTTTGTTCATCGCCGCGCACCTTGGCGTT	180
Qy	181	GATCTTTTGTCCAATGCCGGCGCGATCGTCTCGATATTATGCGCTGGGGTGGCATCGCT	240
Db	181	GATCTTTTGTCCAATGCCGGCGCGATCGTCTCGATATTATGCGCTGGGGTGGCATCGCT	240
Qy	241	TACCTGTTATGTTTGCCTGTCATGCGAGCGAAAGACGCATGACAAACAAGCTGGAAGCG	300
Db	241	TACCTGTTATGTTTGCCTGTCATGCGAGCGAAAGACGCATGACAAACAAGCTGGAAGCG	300
Qy	301	CCACAGATCATTTGAAGAAACAGAACCAACCGTCCCGATGACACGCTTTTGGCGGTTCTG	360
Db	301	CCACAGATCATTTGAAGAAACAGAACCAACCGTCCCGATGACACGCTTTTGGCGGTTCTG	360
Qy	361	GCGGTGGCACTGACACGCGCAACCGGTGCGGGTGGAGGTGAGCGTCGATTAAGCACGG	420
Db	361	GCGGTGGCACTGACACGCGCAACCGGTGCGGGTGGAGGTGAGCGTCGATTAAGCACGG	420
Qy	421	GTTTGGTAAAGCCCATGTTTGATGCAATCTGCTCACTGTTTGAACCCGAATCGCTAT	480
Db	421	GTTTGGTAAAGCCCATGTTTGATGCAATCTGCTCACTGTTTGAACCCGAATCGCTAT	480
Qy	481	TTGACGCGTGTGTGTATTATCGCGCGCGTCCGCGCCAAATACGCGACACCGGACGCTGG	540
Db	481	TTGACGCGTGTGTGTATTATCGCGCGCGTCCGCGCCAAATACGCGACACCGGACGCTGG	540
Qy	541	ATTTTTCGCGCTGGCGGTTTCGGGGCAAGCCTGATCTGGTTCCCGCTGCTGGGGTTTCGGC	600
Db	541	ATTTTTCGCGCTGGCGGTTTCGGGGCAAGCCTGATCTGGTTCCCGCTGCTGGGGTTTCGGC	600
Qy	601	GCAGACGATTTGTACGCGCGCTGTCCAGCGCCCAAGGTGTGGCTCGATCAACCTCGTC	660
Db	601	GCAGACGATTTGTACGCGCGCTGTCCAGCGCCCAAGGTGTGGCTCGATCAACCTCGTC	660
Qy	661	GTGCGAGTTGTGATGACCGCATTGCCCATCAAACTGATTTGATGGGTTAG	711
Db	661	GTGCGAGTTGTGATGACCGCATTGCCCATCAAACTGATTTGATGGGTTAG	711

## RESULT 2

```

US-09-746-660A-51
; Sequence 51, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746, 660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08

```

```

; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patentin Vers. 2.0
; SEQ ID NO 51
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(799)
; OTHER INFORMATION: RXA01394
US-09-746-660A-51

Query Match      100.0%; Score 711; DB 10; Length 822;
Best Local Similarity 100.0%; Pred. No. 4e-225;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  AUGGTGATCATGAAATCTTCAATACAGGTCTGCTTTTGGGGCCAGTCGTTTACTGTGCC 60
DB      92  AUGGTGATCATGAAATCTTCAATACAGGTCTGCTTTTGGGGCCAGTCGTTTACTGTGCC 151

QY      61  ATCGGACCGCAGAATGTACTGCTGATTTAAACAAGGAATTAAGCGCGAAGCACTCATTTGGC 120
DB      152  ATCGGACCGCAGAATGTACTGCTGATTTAAACAAGGAATTAAGCGCGAAGCACTCATTTGGC 211

QY      121  GTTCTTCTGCTGTGTTAAATTTCTGACGCTCTTTTGTTCATCGCGGCACCTCTGGCGGTT 180
DB      212  GTTCTTCTGCTGTGTTAAATTTCTGACGCTCTTTTGTTCATCGCGGCACCTCTGGCGGTT 271

QY      181  GATCTTTTGTCCAATGCGCGCGGATCGTGTGATATTAATGCGCTGGGGTGGCATCGCT 240
DB      272  GATCTTTTGTCCAATGCGCGCGGATCGTGTGATATTAATGCGCTGGGGTGGCATCGCT 331

QY      241  TACCTGTTATGTTTCCGCTCATGCGCAGCGAAGACGCCATGACAAACAAGGTGGAAAGCG 300
DB      332  TACCTGTTATGTTTCCGCTCATGCGCAGCGAAGACGCCATGACAAACAAGGTGGAAAGCG 391

QY      301  CCACAGATCATTTGAAGAAACAGAACCAACCGTCCGATGACACGCTTTTGGCGGGTTGG 360
DB      392  CCACAGATCATTTGAAGAAACAGAACCAACCGTCCGATGACACGCTTTTGGCGGGTTGG 451

QY      361  GCGGTGGCCATGACACGCGCAACGCGGTGCGGGTGGAGGTGAGCGTTCGATAAGCAGCGG 420
DB      452  GCGGTGGCCATGACACGCGCAACCGGGTCCGGGTGGAGGTGAGCGTTCGATAAGCAGCGG 511

QY      421  GTTTGGGTAAAGCCCATGTTGATGCGCAATCGTGTGACTGCTGTTGAACCCGGAATGCCGTA 480
DB      512  GTTTGGGTAAAGCCCATGTTGATGCGCAATCGTGTGACTGCTGTTGAACCCGGAATGCCGTA 571

QY      481  TTGACGCGGTTTGTGTTTATCGCGCGCGTCCGCGCGCAATACGCGCAACCGGACGCTGG 540
DB      572  TTGACGCGGTTTGTGTTTATCGCGCGCGTCCGCGCGCAATACGCGCAACCGGACGCTGG 631

QY      541  ATTTTCGCGCTGCGCGGTTTCGCGCAAGCTGATCTGTTCCGCTGCTGTGGTTTCGCG 600
DB      632  ATTTTCGCGCTGCGCGGTTTCGCGCAAGCTGATCTGTTCCGCTGCTGTGGTTTCGCG 691

QY      601  GCACGACATTTGTCAAGCCCGCTGTTCAGCGCCCAAGAGTGTGGCGCTGGATCAACGTCGTC 660
DB      692  GCACGACATTTGTCAAGCCCGCTGTTCAGCGCCCAAGAGTGTGGCGCTGGATCAACGTCGTC 751

QY      661  GTGCGAGTTGTGATGACCGCAATGCGCAATCAAACTGATGTGATGGGTTAG 711
DB      752  GTGCGAGTTGTGATGACCGCAATGCGCAATCAAACTGATGTGATGGGTTAG 802

```

### RESULT 3

RESUME 3  
US-10-196-232-24  
; Sequence 24, Application US/10196232  
; Publication No. US2003011399A1  
; GENERAL INFORMATION:  
; APPLICANT: YAMAGUCHI, MIKIKO  
; APPLICANT: ITO, HISAO  
; APPLICANT: GUNJI, YOSHIYA

```

; APPLICANT: YASUEDA, HISASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: 225391US0
; CURRENT APPLICATION NUMBER: US/10/196,232
; CURRENT FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: JP 2001-224586
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 2374
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1025)..(1723)
; OTHER INFORMATION:
US-10-196-232-24

Query Match      100.0%; Score 711; DB 14; Length 2374;
Best Local Similarity 100.0%; Pred. No. 6,7e-225;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGTGATCATGAAATCTTCATTACAGCTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 60
DB      1016  ATGTGATCATGAAATCTTCATTACAGCTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 1075

QY      61  ATCGACCGCAGATGTACTGTGATTAAACAAGAAATTAAGCGCGAAGACTCATTTGGC 120
DB      1076  ATCGACCGCAGATGTACTGTGATTAAACAAGAAATTAAGCGCGAAGACTCATTTGGC 1135

QY      121  GTTCTTCTCGTGTGTTTAATTTCTGACGCTCTTTTGTTCATCCCGGCACTTGGGGCTT 180
DB      1136  GTTCTTCTCGTGTGTTTAATTTCTGACGCTCTTTTGTTCATCCCGGCACTTGGGGCTT 1195

QY      181  GATCTTTTGTCCAAATGCGCGCGATCGTCTGCATATTATGCTGCGGTGGCATCGCT 240
DB      1196  GATCTTTTGTCCAAATGCGCGCGATCGTCTGCATATTATGCTGCGGTGGCATCGCT 1255

QY      241  TACTGTGTATGTTTGGCTCATGGCAGGAAAGACGCCATGACAAACAAGGTGGAAGCG 300
DB      1256  TACTGTGTATGTTTGGCTCATGGCAGGAAAGACGCCATGACAAACAAGGTGGAAGCG 1315

QY      301  CCACAGATCATTTGAAGAAAGACAAACCGTGGCCGATGACACGCTTTGGGGCGTTCG 360
DB      1316  CCACAGATCATTTGAAGAAAGACAAACCGTGGCCGATGACACGCTTTGGGGCGTTCG 1375

QY      361  GCGGTGGCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTGCATAAGCAGCG 420
DB      1376  GCGGTGGCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTGCATAAGCAGCG 1435

QY      421  GTTTGGGTAAAGCCATGTTGATGGCAGATCGTGTGACCTGGTTGAAACCGAATGCGTAT 480
DB      1436  GTTTGGGTAAAGCCATGTTGATGGCAGATCGTGTGACCTGGTTGAAACCGAATGCGTAT 1495

QY      481  TTGCAACGCTTTGTGTTTATCGCGCGCTCGCGCGCAATACCGGCAACCGGAGCTGG 540
DB      1496  TTGCAACGCTTTGTGTTTATCGCGCGCTCGCGCGCAATACCGGCAACCGGAGCTGG 1555

QY      541  ATTTTCCGCGCTGGCGCGTTCGCGCAAGCCTGATCTGTTCCCGCTGGTGGGTTTCGGC 600
DB      1556  ATTTTCCGCGCTGGCGCGTTCGCGCAAGCCTGATCTGTTCCCGCTGGTGGGTTTCGGC 1615

QY      601  GCAGACATTTGACGCGCGCTGTCAGCCCAAGAGTGTGGCTGTGATCAACGTCGTC 660
DB      1616  GCAGACATTTGTCAGCGCGCTGTCAGCCCAAGAGTGTGGCTGTGATCAACGTCGTC 1675

QY      661  GTGGCAGTTGTGATGACCGCATTTGGCCATCAACTGATGTTGATGGGTTAG 711
DB      1676  GTGGCAGTTGTGATGACCGCATTTGGCCATCAACTGATGTTGATGGGTTAG 1726

```

RESULT 4

```

US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match      100.0%; Score 711; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 2.4e-223;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGTGATCATGAAATCTTCATTACAGCTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 60
DB      1328953  ATGTGATCATGAAATCTTCATTACAGCTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 1328894

QY      61  ATCGGACCGCAGATGTACTGTGATTAAACAAGAAATTAAGCGCGAAGACTCATTTGGC 120
DB      1328893  ATCGGACCGCAGATGTACTGTGATTAAACAAGAAATTAAGCGCGAAGACTCATTTGGC 1328834

QY      121  GTCCTTCTCGTGTGTTTAATTTCTGACGCTCTTTTGTTCATPCGCGCGCACTTGGGGCTT 180
DB      1328833  GTTCTTCTCGTGTGTTTAATTTCTGACGCTCTTTTGTTCATPCGCGCGCACTTGGGGCTT 1328774

QY      181  GATCTTTTGTCCAAATGCGCGCGATCGTCTGCATATTATGCGCTGGGGTGGCATCGCT 240
DB      1328773  GATCTTTTGTCCAAATGCGCGCGATCGTCTGCATATTATGCGCTGGGGTGGCATCGCT 1328714

QY      241  TACTGTGTATGTTTGGCTCATGGCAGGAAAGACGCCATGACAAACAAGGTGGAAGCG 300
DB      1328713  TACTGTGTATGTTTGGCTCATGGCAGGAAAGACGCCATGACAAACAAGGTGGAAGCG 1328654

QY      301  CCACAGATCATTTGAAGAAAGACAAACCGTGGCCGATGACAGCTTTTGGGGCTTCG 360
DB      1328653  CCACAGATCATTTGAAGAAAGACAAACCGTGGCCGATGACAGCTTTTGGGGCTTCG 1328594

QY      361  GCGGTGGCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTGCATAAGCAGCGG 420
DB      1328593  GCGGTGGCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTGCATAAGCAGCGG 1328534

QY      421  GTTTGGGTAAAGCCATGTTGATGGCAATGTCGTGACCTGTTGTAACCCGGAATGCGTAT 480
DB      1328533  GTTTGGGTAAAGCCATGTTGATGGCAATGTCGTGACCTGTTGTAACCCGGAATGCGTAT 1328474

QY      481  TTGCAACGCTTTGTGTTTATCGCGCGCTCGCGCGCAATACCGGCAACCGGAGCTGG 540
DB      1328473  TTGCAACGCTTTGTGTTTATCGCGCGCTCGCGCGCAATACCGGCAACCGGAGCTGG 1328414

QY      541  ATTTTCCGCGCTGGCGCGTTCGCGCAAGCCTGATCTGCTTCCGCTGGTGGGTTTCGGC 600

```

Db 1328413 ATTTTCGCGCTGCGCGCTTCGCGCAAGCCTGATCTGCTTTCGCGCTTTCGCG 1328354  
 QY 601 GCACGACATTTGACGCGCGCTTCGCGCCCAAGGTGCGCGTGCATCAAGCTGTC 660  
 Db 1328353 GCACGACATTTGACGCGCGCTTCGCGCCCAAGGTGCGCGTGCATCAAGCTGTC 1328294  
 QY 661 GTGGCAGTGTGATGACCGCATTCGCGCATCAAACTGATGTTGATGGGTAG 711  
 Db 1328293 GTGGCAGTGTGATGACCGCATTCGCGCATCAAACTGATGTTGATGGGTAG 1328243

RESULT 5  
 US-09-738-626-3455  
 ; Sequence 3455, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAKO  
 ; APPLICANT: SENO, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OKAZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738, 626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 3455  
 ; LENGTH: 708  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-3455

Query Match 99.6%; Score 708; DB 9; Length 708;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-224;  
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGTGATCATGGAATCTTCAATACAGGTCTGCTTTGGGGGCCAGTCTTTTACTGTC 60  
 Db 1 ATGGTGATCATGGAATCTTCAATACAGGTCTGCTTTGGGGGCCAGTCTTTTACTGTC 60  
 QY 61 ATCGACCGCAGATGCTGCTGATTAACAGGAATTAAGCGGAGGACTCATTTGCG 120  
 Db 61 ATCGACCGCAGATGCTGCTGATTAACAGGAATTAAGCGGAGGACTCATTTGCG 120  
 QY 121 GTTCTTCTCGTGTGTTTAAATTTCTGACGTCTTTTTCATCGCGCGGACCTTTGGCGGT 180  
 Db 121 GTTCTTCTCGTGTGTTTAAATTTCTGACGTCTTTTTCATCGCGCGGACCTTTGGCGGT 180  
 QY 181 GATCTTTTGTCCATGCGCGGATCGTCTGATTAATGCGCTGGGTGGCATCGCT 240  
 Db 181 GATCTTTTGTCCATGCGCGGATCGTCTGATTAATGCGCTGGGTGGCATCGCT 240  
 QY 241 TACCTGTTATGTTTGGTTCATGCGGAAAGCGCCATGACAAACAGGTGGAAGCG 300  
 Db 241 TACCTGTTATGTTTGGTTCATGCGGAAAGCGCCATGACAAACAGGTGGAAGCG 300  
 QY 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTCCCGATGACAGCCCTTTGGCGGTTTCG 360  
 Db 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTCCCGATGACAGCCCTTTGGCGGTTTCG 360

QY 361 GCGGTGGCCACTGACACGCGCAACCGGTCGCGGTGGAGGTGAGCTCGATAAGCAGCG 420  
 Db 361 GCGGTGGCCACTGACACGCGCAACCGGTCGCGGTGGAGGTGAGCTCGATAAGCAGCG 420  
 QY 421 GTTGGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGTTGACCCGCAATCGGTAT 480  
 Db 421 GTTGGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGTTGACCCGCAATCGGTAT 480  
 QY 481 TTGACGCGGTTTGTGTTATCGCGCGCTCGCGCGCAATACGGCGCACACCGGACGTTGG 540  
 Db 481 TTGACGCGGTTTGTGTTATCGCGCGCTCGCGCGCAATACGGCGCACACCGGACGTTGG 540  
 QY 541 ATTTTCGCGCTGCGCGCTTCGCGCAAGCCTGATCTGCTTCCGCTGGTGGGTTCGCGC 600  
 Db 541 ATTTTCGCGCTGCGCGCTTCGCGCAAGCCTGATCTGCTTCCGCTGGTGGGTTCGCGC 600  
 QY 601 GCACGACATTTGACGCGCGCTTCGCGCCCAAGGTGCGCGCTGCGATCAACGTGTC 660  
 Db 601 GCACGACATTTGACGCGCGCTTCGCGCCCAAGGTGCGCGCTGCGATCAACGTGTC 660  
 QY 661 GTGGCAGTGTGATGACCGCATTTGCGCATCAAACTGATGTTGATGGGT 708  
 Db 661 GTGGCAGTGTGATGACCGCATTTGCGCATCAAACTGATGTTGATGGGT 708

RESULT 6  
 US-10-166-142-9  
 ; Sequence 9, Application US/10166142  
 ; Publication No. US20030124687A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GUNJI, YOSHIYA  
 ; APPLICANT: YASUEDA, HISASHI  
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL  
 ; FILE REFERENCE: 223789US  
 ; CURRENT APPLICATION NUMBER: US/10/166,142  
 ; CURRENT FILING DATE: 2002-06-11  
 ; PRIOR APPLICATION NUMBER: JP 2001-177075  
 ; PRIOR FILING DATE: 2001-06-12  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 9  
 ; LENGTH: 712  
 ; TYPE: DNA  
 ; ORGANISM: Brevibacterium lactofermentum  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(375)  
 ; OTHER INFORMATION:  
 US-10-166-142-9

Query Match 98.0%; Score 696.8; DB 14; Length 712;  
 Best Local Similarity 99.6%; Pred. No. 1.9e-220;  
 Matches 709; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 QY 1 ATGGTGATCATGGAATCTTCAATACAGGTCTGCTTTGGGGGCCAGTCTTTTACTGTC 60  
 Db 1 ATGGTGATCATGGAATCTTCAATACAGGTCTGCTTTGGGGGCCAGTCTTTTACTGTC 60  
 QY 61 ATCGACCGCAGAAATGCTGCTGATTAACAGGAATTAAGCGGAGGACTCATTTGCG 120  
 Db 61 ATCGACCGCAGAAATGCTGCTGATTAACAGGAATTAAGCGGAGGACTCATTTGCG 120  
 QY 121 GTTCTTCTCGTGTGTTTAAATTTCTGACGTCTTTTTCATCGCGCGGACCTTTGGCGGT 180  
 Db 121 GTTCTTCTCGTGTGTTTAAATTTCTGACGTCTTTTTCATCGCGCGGACCTTTGGCGGT 180  
 QY 181 GATCTTTTGTCCATGCGCGGATCGTCTGATTAATGCGCTGGGTGGCATCGCT 240  
 Db 181 GATCTTTTGTCCATGCGCGGATCGTCTGATTAATGCGCTGGGTGGCATCGCT 240  
 QY 241 TACCTGTTATGTTTGGTTCATGCGGAAAGCGCCATGACAAACAGGTGGAAGCG 300

241	TACCTGTTATGGTTTGGCTGCTATGCGAGCGAAGACGCCCATGACAAACAAAGGTGGAAAGCG	300
301	CCACAGATCATTTGAAGAAACAAACCAACCGTCGCCGATGACACGCGCTTTGGGCG-GTTC	359
301	CCACAGATCATTTGAAGAAACAAACCAACCGTCGCCGATGACACGCGCTTTGGGCGTGTTC	360
360	GGCGGTGGCCACTGACACCGGCGCAACCGGGTTCGGGTGGAGGTGAGCGTGATAGCAAGCG	419
361	GGCGGTGGCCACTGACACCGGCGCAACCGGGTTCGGGTGGAGGTGAGCGTGATAGCAAGCG	420
420	GGTTTGGGTAAAGCCCATGTCATGGCAATCGTCTGACCTGGTTGAACCCGAATGCGTA	479
421	GGTTTGGGTGAAGCCCATGTCATGGCAATCGTCTGACCTGGTTGAACCCGAATGCGTA	480
480	TTTGGACCGGTTGTGTATTATCGGCGCGCTCGCGCGCAATA	539
481	TTTGGACCGGTTGTGTATTATCGGCGCGCTCGCGCGCAATA	540
540	GATTTTCCGCGCTGGCGCGCTCGGGCAAGCCTGATCTGGTTCCCGCTGGTGGGTTTCGG	599
541	GATTTTCCGCGCTGGCGCGCTCGGGCAAGCCTGATCTGGTTCCCGCTGGTGGGTTTCGG	600
600	CGCAGCAGCATTTGTCAAGCCCGCTGTCCAGCCCAAGGTGTGGCGCTTGGATCAACGTCGT	659
601	CGCAGCAGCATTTGTCAAGCCCGCTGTCCAGCCCAAGGTGTGGCGCTTGGATCAACGTCGT	660
660	CGTGGAGTTGTGATGACCGCATTTGGCGCATCAACATGATTTGATGGGTTAG	711
661	CGTGGAGTTGTGATGACCGCATTTGGCGCATCAACATGATTTGATGGGTTAG	712

```

RESULT 7
US-10-156-761-1109
/ Sequence 1109, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 1109
/ LENGTH: 609
/ TYPE: DNA
/ ORGANISM: Streptomyces avermitilis
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(609)
US-10-156-761-1109

```

Qy	505	GGCGTCGGCGCGCAATACCGCGACACCGGACGGTGGATTTTCGCGCGTCGGCGCGTTCGGG	564
Db	406	GGCTCGATCGCGGCGGACCGCGGCTCACTGCGCTGGACCTTTCGCTCTCGGGCGCGGCGCTC	465
Qy	565	GCAAGCGCTGATCGCTTTCGCGCTGGTGGGTTTCGGGCGGACGACGATGTCAGGCGCGCTG	624
Db	466	GCAGCGCTGTGCTGGTTCGCGCGCTCGGCTTCGGCTCACGGCTGCTCGGCGGCTTCTCTG	525
Qy	625	TCAGCGCCCAAGGTGGCGGCTGGGATCAACGTCGTCGTGGCGAGTTGTGATGACCGCGATTG	684
Db	526	GCAGCGCGCTCGGATGCGGCGTGTGAGCGCCCTGTCGCGCGCGACGATGCTCACGATG	585
Qy	585	GCCATCAAACTGAT	698
Db	586	GGCGGCAACCTTGAT	599

RESULT 8  
 US-10-156-761-1/c  
 ; Sequence 1, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: HATTORI, MASAHIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 15109  
 ; SEQ ID NO 1  
 ; LENGTH: 9025608  
 ; TYPE: DNA  
 ; ORGANISM: Streptomyces avermitilis  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4187715)  
 ; OTHER INFORMATION: a, t, c, g, other or unknown  
 US-10-156-761-1

Query Match	8.8%;	Score 62.8;	DB 14;	Length 9025608;	
Best Local Similarity	52.2%;	Pred. No. 5.5e-08;			
Matches 164;	Conservative	0;	Mismatches 147;	Indels 3;	Gaps 1;
QY	385	CGGGTGGGGTGGAGGTGAGCGGTGCATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATG	444		
Db	1406101	CGGGCGGGCTCTCGGGCGCGAAGGGCGCGCGAGGGCTCCCGACAGCGCGCTACTACCC	1406042		
QY	445	GCAATCGTGTCTGACTGGTTGAACCCGAATGCGTATTTGACGCGTTTGTATTATCGCG	504		
Db	1406041	TGCTTGGCCCTCACTTGGCTCAACCCGACGCTCACTTCGACACC---GTCTTCCTCTCT	1405985		
QY	505	GGCGTGGCGCGCAATACGGGACACCGGACGGTGGATTTTGGCGCGCTGGCGCGCTTCGG	564		
Db	1405984	GGCTCGATTCGCGCGCCGACGCGCGGTCACTGCGCTGGACCTTCGGTCTCGCGGGCGCTTC	1405925		
QY	565	GCAAGCTGATCTGGTTCCCGCTGGTGGGTTTCGGCGCAGCAGCATTTGTCAAGCCCGCTG	624		
Db	1405924	GGAGCGCTGTCTGGTTGCGCGCCCTCGGCTTCGCGCTCACGCGTCTCGCGCGCTTCCTG	1405865		
QY	625	TCACGCCCAAGGTGTGGCGCTGATCAACGTCGCTGTGGCAGTTGTGATGACCGCATTG	684		
Db	1405864	GGCGCCGTGTGGCATTGGCGGGGTCTGGAGAGCCCTTGTGTGGCGGACGATGCTCAGCATG	1405805		
QY	685	GCCATCAAACTGAT	698		



```

Db 1405804 GCGCCACCTGAT 1405791
|||||
RESULT 9
US-09-894-844-11
; Sequence 11, Application US/09894844
; Patent No. US20020176873A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; FILE REFERENCE: STAN102CON
; CURRENT APPLICATION NUMBER: US/09/894,844
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-09-894-844-11

Query Match 8.0%; Score 57; DB 9; Length 597;
Best Local Similarity 53.1%; Pred. No. 3.9e-08;
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 431 AGCCATGTTGATGGCAATCGTGTGACCTGGTTGAACCCGATGCGTATTTGGACGGT 490
Db 326 AGGTCTGTGACCTGTGGCGCATTCAGTTCCTCAACCCACAGTCTACCTCGACACG 385
QY 491 TTGTGTTTATCGCGCGCTGCGCGCAATACGCGGACACCGGATGCGTATTTGGCGG 550
Db 386 TCGTGTGCTAGCGCGCTGGCCAAACGACGACAGCA--CCAGCGCTGGCTGTTGGCC 442
QY 551 CTGCGCGCTTGGCGGCAAGCCTGATCTGTTCCCGTGGTGGTTTCGGCGCAGCAT 610
Db 443 TCGCGCGCTTACAGCCAGTGGGTATGTTGCGCCACCTCGGGTTCGGAGCGCGCGT 502
QY 611 TGTACGCGCGCTGTCAGCCCAAGGTGGCGCTGGATCAAGTCTGCTGGCGAGTTG 670
Db 503 TCGCGCGCTGTTCAACACCCCGCTCGTGGAGATCTCTGACGGCTGATCGCGTCA 562
QY 671 TGATGACCGCATTTGGCCATCAAACTGATGTTGA 703
Db 563 TGATGTTGCGCTGGGAATCTCGTGACCGTGA 595

RESULT 10
US-10-080-170-648
; Sequence 11, Application US/10388902
; Publication No. US20040018574A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; FILE REFERENCE: STAN102CON
; CURRENT APPLICATION NUMBER: US/10/388,902
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US/09/894,844
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25

```

```

; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-10-388-902-11

Query Match 8.0%; Score 57; DB 15; Length 597;
Best Local Similarity 53.1%; Pred. No. 3.9e-08;
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 431 AGCCATGTTGATGGCAATCGTGTGACCTGGTTGAACCCGATGCGTATTTGGACGGT 490
Db 326 AGGTCTGTGACCTGTGGCGCATTCAGTTCCTCAACCCACAGTCTACCTCGACACG 385
QY 491 TTGTGTTTATCGCGCGCTGCGCGCAATACGCGGACACCGGATGCGTATTTGGCGG 550
Db 386 TCGTGTGCTAGCGCGCTGGCCAAACGACGACAGCA--CCAGCGCTGGCTGTTGGCC 442
QY 551 CTGCGCGCTTGGCGGCAAGCCTGATCTGTTCCCGTGGTGGTTTCGGCGCAGCAT 610
Db 443 TCGCGCGCTTACAGCCAGTGGGTATGTTGCGCCACCTCGGGTTCGGAGCGCGCGT 502
QY 611 TGTACGCGCGCTGTCAGCCCAAGGTGGCGCTGGATCAAGTCTGCTGGCGAGTTG 670
Db 503 TCGCGCGCTGTTCAACACCCCGCTCGTGGAGATCTCTGACGGCTGATCGCGTCA 562
QY 671 TGATGACCGCATTTGGCCATCAAACTGATGTTGA 703
Db 563 TGATGTTGCGCTGGGAATCTCGTGACCGTGA 595

RESULT 11
US-10-080-170-648
; Sequence 648, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 648
; LENGTH: 86114
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-648

Query Match 7.9%; Score 56.2; DB 14; Length 86114;
Best Local Similarity 53.2%; Pred. No. 8.5e-07;
Matches 143; Conservative 0; Mismatches 123; Indels 3; Gaps 1;

QY 437 TGTGATGGCAATCGTGTGACCTGGTTGAACCCGATGCGTATTTGGACGGTGTGTGT 496
Db 64559 TGCAAAATGTGCTGGTGGTGACCTTTCTCAACCCACAGCTCTATCTGGACACTGTGTGT 64618
QY 497 TTATCGCGCGCTGCGCGCAATACGCGGACACACCGGATGCGTATTTGGCGGTGGCG 556
Db 64619 TGATCGGTGCTGCGCATAGAGGATCAGA---TCTGGGTGGTTTTTCGAGCGCGGTG 64675
QY 557 CGTTCGCGCGCAAGCCTGATCTGTTCCCGTGGTGGTTTCGGCGCAGCAGCATTTGTCA 616
Db 64676 CCTGGCGCGCGCAGCGTGTATGTTTGGGATTTAGCGCGGCGCGCTACAGC 64735

```

QY 617 GCGCGTGTCCAGCCCAAGGTGGCGTGGATCAAGTCGTGGCGGTGGATGA 676  
 Db 64736 CATTTTCGCACTCACTGCTTGGCGCATTTCTTGATGCGTGGTTCGCGGACGATGA 64795  
 QY 677 CCGCATTTGGCCATCAAACTGATGTTGATG 705  
 Db 64796 TTGGGGTGGCGCGTGTGTGTCTGTCTCATCG 64824

RESULT 12

US-10-156-761-5949  
 ; Sequence 5949, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: HATTORI, MASAHIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 15109  
 ; SEQ ID NO 5949  
 ; LENGTH: 3000  
 ; TYPE: DNA  
 ; ORGANISM: Streptomyces avermitilis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(3000)  
 US-10-156-761-5949

Query Match 5.9%; Score 42.2; DB 14; Length 3000;  
 Best Local Similarity 51.9%; Pred. No. 0.0072;  
 Matches 123; Conservative 0; Mismatches 108; Indels 6; Gaps 1;  
 QY 369 CACTGACACGCGCAACCGGCTGGAGGTGGAGTGCATGATGAGCAGCGGTTGGGT 428  
 Db 57 CACCGATCGTGGCGCGCGAGCGGTGCTGGAGAGCGCGAGCTGGAGCCCGTCAAGGC 116  
 QY 429 AAAGCCCATGTTGATGGCAATCGTGTGACCTGTTGAACCCGAAATGCGTATTTGGACGC 488  
 Db 117 CGACCCCGTGTCTGGAAGCCCTG-----GGCGGACGCGCGATCCCGATCTGGCGCT 170  
 QY 489 GTTTGTGTTTATCGCGCGCTGGCGCGCAATACGCGACACCGGACGGTGGATTTTCGC 548  
 Db 171 GCTCGACTGCTCCCGCTGGTGGAGGCGAGAGCAGACACCGCCCGCGGAGCTCT 230  
 QY 549 CGCTGGCGGTTTCGCGGCAAGCTGATCTGTTCCCGCTGGTGGGTTTCGCGCGCAGC 605  
 Db 231 CGACACGCTGATCGCGGCGCAACCCCTGCGGACCGGCTCTCTCGGCGTCTCTCGGTGC 287

RESULT 13

US-10-156-761-1  
 ; Sequence 1, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: HATTORI, MASAHIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 15109  
 ; SEQ ID NO 1  
 ; LENGTH: 9025608  
 ; TYPE: DNA  
 ; ORGANISM: Streptomyces avermitilis  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4187715)  
 ; OTHER INFORMATION: a, t, c, g, other or unknown  
 US-10-156-761-1  
 Query Match 5.9%; Score 42.2; DB 14; Length 9025608;  
 Best Local Similarity 51.9%; Pred. No. 0.38;  
 Matches 123; Conservative 0; Mismatches 108; Indels 6; Gaps 1;  
 QY 369 CACTGACACGCGCAACCGGCTGGAGGTGGAGTGCATGATGAGCAGCGGTTGGGT 428  
 Db 7187526 CACCGATCGTGGCGCGCGAGCGGTGCTGGAGAGCGCGAGCTGGAGCCCGTCAAGGC 7187585  
 QY 429 AAAGCCCATGTTGATGGCAATCGTGTGACCTGTTGAACCCGAAATGCGTATTTGGACGC 488  
 Db 7187586 CGACCCCGTGTCTGGAAGCCCTG-----GGCGGACGCGCGATCCCGATCTGGCGCT 7187639  
 QY 489 GTTTGTGTTTATCGCGCGCTGGCGCGCAATACGCGACACCGGACGGTGGATTTTCGC 548  
 Db 7187640 GCTCGACTGCTCCCGCTGGTGGAGGCGAGGACGACACACCGCCCGCGGAGCTCT 7187699  
 QY 549 CGCTGGCGGTTTCGCGGCAAGCTGATCTGTTCCCGCTGGTGGGTTTCGCGCGCAGC 605  
 Db 7187700 CGACACGCTGATCGCGGCGCAACCCCTGCGGACCGGCTCTCTCGGCGTCTCTCGGTGC 7187756

RESULT 14

US-10-369-493-31368  
 ; Sequence 31368, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 31368  
 ; LENGTH: 1392  
 ; TYPE: DNA  
 ; ORGANISM: Burkholderia cepacia  
 US-10-369-493-31368  
 Query Match 5.7%; Score 40.4; DB 15; Length 1392;  
 Best Local Similarity 51.1%; Pred. No. 0.019;  
 Matches 95; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
 QY 417 GCGGTTTGGTAAAGCCCATGTTGATGCAATCGTGTGACCTGTTGAACCGGAATGC 476  
 Db 793 GCGGTGGCGGTGGCGGGAAGTTTCAGACCTCGGGGCAAGATTGTTCGCGGCGAATCG 842  
 QY 477 GTATTGGACGCGTTTGTGTTTATCGCGCGGCTCGGCGCGCAATACGCGGACACCGGAGC 536

Db 843 CATTTCCTCGCGCGCGCTGTACGAAGCGTTCGTCGCCCAATACAGCGAGCGCGTCAA 902  
 QY 537 GTGGATTTTCGGCGCGCTTCGCGCAAGCTGATCTGGTTCCCGCTGGTGGTTT 596  
 Db 903 GCGCTGAAAGTTCGGCGCTGCTTCGAAGCGGACGTGGACGTCGCGCCCGTTGATGATCA 962  
 QY 597 CGCGCG 602  
 Db 963 AGCCGC 968

RESULT 15  
 US-10-369-493-28610  
 ; Sequence 28610, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; PRIOR FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 28610  
 ; LENGTH: 1443  
 ; TYPE: DNA  
 ; ORGANISM: Burkholderia fungorum  
 US-10-369-493-28610

Query Match 5.7%; Score 40.4; DB 15; Length 1443;  
 Best Local Similarity 51.1%; Pred. No. 0.02;  
 Matches 95; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 417 GCGGTTTGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGTTGAACCCGAATGC 476  
 Db 825 GCGCGTGGCGTTCGCGCGAAGTTTCAGACCTCGGGGACAGGATTGTTGCGCGCGAATCG 884  
 QY 477 GTATTGACGCGTTTGTGTTATCGCGGCGTTCGCGGCGCAATACGGGACACCGGACG 536  
 Db 885 CATTTCGTCGCGGCGCGCTGTACGAAGCGTTCGTCGCCCAATACAGCGCGCGTCAA 944  
 QY 537 GTGGATTTTCGCGCGCTGCGCGCGCTTCGCGCAAGCTGATCTGTTCCCGCTGGTGGTTT 596  
 Db 945 GCGCTGAAAGTTCGCGCTGCGTTTCGAAGCGACGTGGACGTCGCGCCCGTTGATGATCA 1004  
 QY 597 CGCGCG 602  
 Db 1005 AGCCGC 1010

Search completed: March 17, 2004, 09:46:58  
 Job time : 282.891 secs

100

C	1	60.4	8.5	830	29	CNS01MQH	ALJ51258 Anopheles
C	2	56.8	8.0	1034	28	BZ554294	BZ554294 pac81-60
C	3	54.2	7.6	897	28	BZ560550	BZ560550 pac82-164
C	4	51.4	7.2	1620	28	BZ568946	BZ568946 pac82-164

## ALIGNMENTS

COMMENT

This clone is from an A

This clone is from an A

Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.

# FEATURES

source  
1. .830  
Location/Qualifiers  
/organism="Anopheles gambiae"  
/mol\_type="genomic DNA"  
/strain="PBST"  
/db\_xref="taxon:7165"  
/clone="22824"  
/clone\_lib="NotreDamel"  
/note="end : T7"

## ORIGIN

Query Match 8.5%; Score 60.4; DB 29; Length 830;  
Best Local Similarity 50.7%; Pred. No. 9.5e-05;  
Matches 142; Conservative 1; Mismatches 137; Indels 0; Gaps 0;  
QY 23 TTACAGCTGCTTTTGGGGCCAGCTTTTACTGTCATCGGACCGGCAATGATCGG 82  
Db 597 ATCAAGGGCTTGCCCTTAGGTGCAATTTTATTCTTCCCTCGGCCGCAAAATCGTTG 538  
QY 83 TGATTAAACAAGGAAATTAAGCGCAAGGACTCATTTGGGTTCTTCTCGTGTGTTTATT 142  
Db 537 TGATGATCAGGGCAATCGCGCCAGTACCATCTGATGATGCTGTATGCGGGTAA 478  
QY 143 CTGACGCTTTTGTGTCATCGCGCACTTGGGCGTTGATCTTTTGTCCCAATCGCGCG 202  
Db 477 GTGATTTGCTGCTGATCTGTCTCGGGATTTTGGCGGCGGCTGCTGATGAGTCTC 418  
QY 203 CGATGCTGCTCGCATATTATGGCTGGGTCGATCGCTTACCTGTTAGTTTCCGTC 262  
Db 417 CGTGCTGCTGGCGTTGCTACCTGGGGCGGCTGGGCTTCTGCTGTGGTACGGATTG 358  
QY 263 TGGCAGCGGAAGACGCCATGACAAACAAGGTGGAAGGCC 302  
Db 357 GCGGCTGAAAGCGGATGAGCAGTAACCTCGAACTGGC 318

# RESULT 2

BZ554294 1034 bp DNA linear GSS 17-DEC-2002  
LOCUS  
DEFINITION  
pacsl-60\_4617.x1 pacsl-60 Pseudomonas aeruginosa genomic clone  
pacsl-60\_4617, genomic survey sequence.  
ACCESSION  
BZ554294  
VERSION  
BZ554294.1 GI:27161466  
KEYWORDS  
GSS.  
SOURCE  
Pseudomonas aeruginosa  
ORGANISM  
Pseudomonas aeruginosa  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
Pseudomonadaceae; Pseudomonas.

# REFERENCE

1 (bases 1 to 1034)  
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,  
Burns, J.L., Kaul, R. and Olsen, M.V.  
Whole-Genome-Sequence variation among multiple isolates of  
Pseudomonas aeruginosa library  
J. Bacteriol. (2002) In press  
Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.

# FEATURES

source  
1. .1034  
Location/Qualifiers  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="1-60"  
/db\_xref="taxon:287"  
/clone="pacsl-60\_4617"  
/clone\_lib="pacsl-60"  
/note="Clinical isolate 1-60 Whole genomic shotgun"

# ORIGIN

Query Match 7.6%; Score 54.2; DB 28; Length 897;  
Best Local Similarity 52.8%; Pred. No. 0.0041;  
Matches 143; Conservative 0; Mismatches 120; Indels 8; Gaps 1;  
QY 440 TGATGGCATCTGCTGACCTGGTTGAAACCCGATGCTATTGGACCGCTTGTGTTA 499  
Db 552 TCGGGGCACTGGCGTTACCTGGCTTAAACCCCAACGCTATTTCGATACCGTCTTCTAC 493

# library."

## ORIGIN

Query Match 8.0%; Score 56.8; DB 28; Length 1034;  
Best Local Similarity 50.4%; Pred. No. 0.0009;  
Matches 139; Conservative 0; Mismatches 137; Indels 0; Gaps 0;  
QY 7 ATCATGGAATCTTCATTACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTCAATCGGA 66  
Db 82 AGTTTTCAGAGCTATCTCAACGGCATCTTGTTGGTGGCGGCTGTATCATCGCCATCGGC 141  
QY 67 CCGCAGATGCTACTGCTGATTAAACAAGGAAATTAAGCGCAAGGACTCATTTGGGTTCTT 126  
Db 142 GCGCAGATGCTTGTCTGCGCAGAGCTGCGCGGAGCATCACTCTCGGTAGCC 201  
QY 127 CTCGTGTGTTAATTTCTGAGTCTTTTGTTCATCGCGGACCTTGGGCGTTGATTT 186  
Db 202 GCGCTCTGCTGTTCTGCGAGCGGCTGCTGTCAAGCTCGGCGTGTTCGCGCTGCGCAAG 261  
QY 187 TTGTCAATGCGCGCGCATGCTGCTGATATTATATGCGCTGGGTTGGCATCGCTTACTG 246  
Db 262 CTGCTGCTGAAACCCGAGCTGCTGCGCATGCGCCCTGCGGCGGAATCGCCTTCTG 321  
QY 247 TTATGTTTCCCTCATGCGCAGCGAAGAGCGCATG 282  
Db 322 ACCTGGTACGGCTCAAGGCGCTCTTCGCGGCTTG 357

## RESULT 3

BZ560550/c 897 bp DNA linear GSS 17-DEC-2002  
LOCUS  
DEFINITION  
pacsl-164\_2569.x1 pacsl-164 Pseudomonas aeruginosa genomic clone  
pacsl-164\_2569, genomic survey sequence.  
ACCESSION  
BZ560550  
VERSION  
BZ560550.1 GI:27178749  
KEYWORDS  
GSS.  
SOURCE  
Pseudomonas aeruginosa  
ORGANISM  
Pseudomonas aeruginosa  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
Pseudomonadaceae; Pseudomonas.

# REFERENCE

1 (bases 1 to 897)  
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,  
Burns, J.L., Kaul, R. and Olsen, M.V.  
Whole-Genome-Sequence variation among multiple isolates of  
Pseudomonas aeruginosa library  
J. Bacteriol. (2002) In press  
Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.

# FEATURES

source  
1. .897  
Location/Qualifiers  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="2-164"  
/db\_xref="taxon:287"  
/clone="pacsl-164\_2569"  
/clone\_lib="pacsl-164"  
/note="Clinical isolate 2-164 Whole genomic shotgun"

# ORIGIN

Query Match 7.6%; Score 54.2; DB 28; Length 897;  
Best Local Similarity 52.8%; Pred. No. 0.0041;  
Matches 143; Conservative 0; Mismatches 120; Indels 8; Gaps 1;  
QY 440 TGATGGCATCTGCTGACCTGGTTGAAACCCGATGCTATTGGACCGCTTGTGTTA 499  
Db 552 TCGGGGCACTGGCGTTACCTGGCTTAAACCCCAACGCTATTTCGATACCGTCTTCTAC 493

```

QY 500 TCGCGCGCTCGGCGCAATACGCGCACACCGGATTTTCGGCGTGGCGGT 559
Db 492 GGTTCGCTCGGCGCGCAGCAAGCGCG-----CGGGCGNTATGCCCTCGCGCGG 441
QY 560 TCGCGGCAACCTGATCTGTTCCGCTGGTGGTTTCGCGCAGCAGCATGTACGCC 619
Db 440 CAAAGCGCTCGCTGATGTGTTTCGCGCTCGCCCTCGCGCGGCGATGGCTGGCGCCCT 381
QY 620 CGCTGTCCAGCCCCCAAGGTGGCGGCTGGATCAAGCTGCTGGCAGTTGTGATGACCG 679
Db 380 GCTTGGCGCGCGCGCCACCTGGCGGCTGCTGACCTGATGTGGCGGCCATGATGCTGG 321
QY 680 CATTGGCCATCAAACTGATGTGATGGTTA 710
Db 320 GCATGGCGCGCAACTGCTGTTCCGGGATA 290

RESULT 4
BZ568946/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 1620)
AUTHORS
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL
J. Bacteriol. (2002) In press
COMMENT
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..1620
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."

FEATURES
source
Query Match 7.2%; Score 51.4; DB 28; Length 1620;
Best Local Similarity 54.3%; Pred. No. 0.027;
Matches 100; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 527 ACACCGGAGGTGGATTTTCGCGCTGGCGGTTTCGCGCAGCCTGATCTGCTCCGC 586
Db 618 AGNCCGCGGCGGCATATGCCCTCGCGCGGCACGCGCTCGATGATGTGTTCTCG 559
QY 587 TGGTGGGTTTCGCGCAGCAGCATTTGTCACGCGCGCTGTCCAGCCCCCAAGGTGTGCGCT 646
Db 558 CCCTCGCCCTCGCGCGCATGCTGCCCCCTGCTGGCGCGGCCCGCNCACCTGCGGCC 499
QY 647 GATCAACCTGCTGTCGATGTGTGATGACCGCATTTGGCCATCAAACTGATGTGATGG 706
Db 498 TGCTGACATGATGTGGCGGCATGATGCTGGGCGATGGCGCGCACTGCTGTTCGGG 439
QY 707 GTTA 710

```

```

Db 438 GATA 435

RESULT 5
BZ556963/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 833)
AUTHORS
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL
J. Bacteriol. (2002) In press
COMMENT
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..833
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone_lib="pacs1-60"
/note="clinical isolate 1-60 Whole genomic shotgun
library."

ORIGIN
Query Match 6.8%; Score 48.6; DB 28; Length 833;
Best Local Similarity 50.6%; Pred. No. 0.12;
Matches 117; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 52 TTATGTCCTATCGGACCCGACATGCTACTGGTGATTAACAGAGATTAAAGCGCAAGGA 111
Db 446 TCATCGGCCCATCGCGCGCGCAAGATGCTTCGTCCTCGCGCAGAGCCTCGCGCGAGCAT 387
QY 112 CTCATTGCGGTTCTTCTCGTGTGTTTAAATTTCTGACGTCCTTTTGTTCATCGCGCGCAC 171
Db 386 CACCTCTCGGTAGCGCGCTCTGCGTGTTCGCGACAGGCTGCTGAGCTCGGCTG 327
QY 172 TTGGGCGTTGATCTTTTGTCCAATGCGCGCGCATCGGTGCTCGATATTATGCGTGGGGT 231
Db 326 TTCGCGCTTGCCCAAGCTGCTCTG3AAACCGACGCTGCTGCCCATCGCCGCTG3GGC 267
QY 232 GGCATGCTTACTCTTATGTTTTCGCTGATGCGAGGAGAAAGCCCATG 282
Db 266 GGGATCGCTTCTTCTGACCTGACCTGACGGGCTCAAGGGCTGCTCGCGCGTTG 216

RESULT 6
CNS0091P
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACK19D16 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL053013
VERSION
AL053013.1 GI:4934461
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)

```

ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 925)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - FRANCE (E-mail : [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr))  
- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Goosawa and Aaron Mamoser in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's 21 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
source  
1..925  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR19D16"  
/clone\_lib="RPCI-98"  
/note="end : TET3"

ORIGIN  
Query Match 6.8%; Score 48.6; DB 29; Length 925;  
Best Local Similarity 13.8%; Pred. No. 0.12;  
Matches 49; Conservative 157; Mismatches 149; Indels 0; Gaps 0;  
QY 325 CCAACGTCGCGATGACAGCGCTTGGCGGTTCGGCGTGGCCATGACAGCGCAAC 384  
DB 571 SC 630  
QY 385 CGCGTCGCGGTGACGATGACGATGACGATGACGATGACGATGACGATGACGATG 444  
DB 631 CSSSSSCSS 690  
QY 445 GCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 504  
DB 691 STGTSSTSS 750  
QY 505 GCGTCGCGCGCAATACGCGACACGCGATGATTTTCGCGCTGCGCGCTTCGCG 564  
DB 751 SSYSSTSCCTCCSS 810  
QY 565 GCAAGCCTGATCTGCTTCCCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTC 624  
DB 811 MCTCCSTYMBCTSTSCGSSSSSSSGKGVTKCGCGCGSSSTNGMTSSSACSSSS 870  
QY 625 TCAGCCCAAGGTGTCGCGCTGATCAACGTCGTCGTCGTCGTCGTCGTCGTCGTC 679  
DB 871 SSSSVSSSSKSS 925

RESULT 7  
BX393687  
LOCUS BX393687 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
DEFINITION cDNA clone CSODC001YG07 5-PRIME, mRNA sequence.  
ACCESSION BX393687  
VERSION BX393687.1 GI:30624032  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 852)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - FRANCE  
Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of Invitrogen. Contact: Feng Liang Email : [liang@lifetech.com](mailto:liang@lifetech.com) URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODC001AD04QP1.

FEATURES  
Location/Qualifiers  
1..852  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODC001YG07"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="First strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 6.7%; Score 47.6; DB 13; Length 852;  
Best Local Similarity 17.2%; Pred. No. 0.22;  
Matches 69; Conservative 152; Mismatches 178; Indels 3; Gaps 1;  
QY 23 TTACAGTCTGCTTTGGGGGCGAGTCTTTTACTGTCATCGACCGCAGAAATGACTGG 82  
DB 381 TTTCKGKTK 440  
QY 83 TGATTAAACAGGAATTAAGCGCGAAGGACTCAATTCGGGTCTCTCTGCTGTGTTAATT 142  
DB 441 KKK 500  
QY 143 CTGACGCTTTTGTTCATCGCGGACCTTGGCGCTTGTATCTTTTGTCCAAATCGCGC 202  
DB 501 KDKKK 560  
QY 203 CGATCGTCTGATATATATGCGTGGCGTGGCATGCTTACCTGTATGTTTTCGCTCA 262  
DB 561 CTTTCKKK 620  
QY 263 ---TGGCAGCGAAGACGCGCATGACAAACAAAGGTGGAAGCGCCACAGATCAATTAAGAAA 319  
DB 621 KTTTCKKK 680  
QY 320 CAGAACCAACGCTGCCGATGACAGCCCTTTTGGCGGTTTCGGCGGTTCGCCACTGACAGC 379  
DB 681 CGGNNVVVVVMMGMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM 740  
QY 380 GCAACCGGTCGCGTGGAGTTCAGCGTTCGATTAAGCAGAGCGGG 421  
DB 741 MMMMMMMGGGMMGGGMMGGGMMGGGMMGGGMMGGGMMGGGMMGGGMMGGG 782

RESULT 8  
CA830247/c  
LOCUS CA830247 614 bp mRNA linear EST 12-DEC-2002  
DEFINITION 1117004E09.y1 1117 - Unigene V from Maize Genome Project Zea mays  
ACCESSION cDNA, mRNA sequence.  
VERSION CA830247  
KEYWORDS CA830247.1 GI:26558012  
SOURCE EST.  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;



Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.

# REFERENCE

1 (bases 1 to 614)

# AUTHORS

Walbot, V.

# JOURNAL

Maize ESTs from various cDNA libraries sequenced at Stanford

# COMMENT

University

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 1117004 row: E column: 09.

# FEATURES

Location/Qualifiers

1..614

/organism="Zea mays"

/mol\_type="mRNA"

/db\_xref="dbEST:946110C04.y1"

/clone\_lib="1117 - Unigene V from Maize Genome Project"

/note="This library represents the unique genes found in the fifth round of EST sequencing at Stanford University for the maize genome project. Sequences are present from library 946. Contigs were assembled using ZmDBAssembler and 2 representatives from each contig were selected for the Unigene set. All singlets were also selected."

# ORIGIN

Query Match 6.5%; Score 46; DB 14; Length 614;

Best Local Similarity 46.3%; Pred. No. 0.51;

Matches 151; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 349 TTGGGGGTTTCGGCGTGGCCATCTACACCGCCACCGGTCGGGTGGAGTGCAGCTC 408

Db 332 TTGGCCCTGACCGCGGAGAGTCCGAGATGCGCCCTTGAGAGCTTCCGGTTATGGTC 273

QY 409 GATAAGCAGCGGTTTGGGTAAGCCCATGTTGATGCGATCTGACCTGCTGTAAC 468

Db 272 CTCTCGTAGTAGACAGGTGGAGCCGAAGTGGACGTAGCAGGTGCGCGGAGTGGATC 213

QY 469 CCGAATCGTATTGACGCGTTTGTGTTTATCGCGCGCGCTCGCGCGCAATACGGCGAC 528

Db 212 TCGAAGCGCGCGCTGCTGCGAAGCTGTAGGAGCGACGAGTCCGGGATGAGCCCG 153

QY 529 ACCGAGCGTGTGATTTTCGCGCTGGCGGTTTCGGCGCAAGCTGATCTGCTCCGCTG 588

Db 152 TTCGGAGGCGCGTACTTGGGAGTAGTGTTCGCGCGCGCTGAGCGCGTGCTGAC 93

QY 589 GTGGGTTTCGGCGCAGCAGCATTTGTACGCGCGCTGTCCAGCCCCCAAGGTGTGGCGTGG 648

Db 92 GCGGTGACCGCTGCG 33

QY 649 ATCAACTGCTGTGGCGAGTTTGAT 674

Db 32 GACATTCCCGTTGGGATATTGGGAT 7

RESULT 9

CD881613

LOCUS F1.103123F010329 F1 Triticum aestivum cDNA clone F1103L23, mRNA

DEFINITION 720 bp mRNA linear EST 14-JUL-2003

sequence.

ACCESSION CD881613

VERSION CD881613.1 GI:32641286

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Triticum.

1 (bases 1 to 720)

# AUTHORS

Genoplatte.

Genoplatte, a major partnership french program in plant genomics

Unpublished (2003)

Contact: Genoplatte

Genoplatte

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french

plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>

and <http://genoplatte-info.infobiogen.fr>).

FEATURES

Location/Qualifiers

1..720

/organism="Triticum aestivum"

/mol\_type="mRNA"

/cultivar="recital"

/db\_xref="taxon:4565"

/clone="F1103L23"

/tissue\_type="leaf one"

/clone\_lib="F1"

ORIGIN

Query Match 6.5%; Score 46; DB 14; Length 720;

Best Local Similarity 48.1%; Pred. No. 0.54;

Matches 130; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 405 COTCGATAAGCAGCGGTTTGGGTAAGCCCATGTTGATGCAATCTGCTGACCTGTT 464

Db 251 COTGGAGAGAGCTGGTGGCGCTCAAGCCCAAGAGCATGACTTCGCCGCGCGCG 310

QY 465 GAACCCGAATGCGTATTGAGAGCGTTTGTGTTTATCGGCGGCTGGCGCGCAATACGG 524

Db 311 TCTGCGCTCGCCATCGAGAGCGCCCATGAGGCGCTCGAGAGCGCGGCTTCTCGCGCG 370

QY 525 CGACACCGAGCGGTGATTTTCGCGCGTGGCGGTTTCGCGCAAGCTCATCTGTTCCC 584

Db 371 CAAGTCCATCTCTGCTCGCGCGCGCGCGGAGTCTCGGACCTCTGTCATCCAGCTGGC 430

QY 585 GCTGTGGTGGTTTCGCGCAGCAGCATTTGTCAGCCCGCTGTCCAGCCCCCAAGGTGTGG 644

Db 431 GAAGCAAGTTTACGGCGCTGCGAGTGGCGGCCAGCCCGGAGCTGGAGCT 490

QY 645 CTGATCAAGCTGCTGCTGGCGAGTTTGAT 674

Db 491 COTGAAAGCCTGGGAGCGCGCTGGCCAT 520

RESULT 10

CA829022

LOCUS 1114036D07.y2 1114 - Unigene IV from Maize Genome Project Zea mays

DEFINITION 620 bp mRNA linear EST 11-DEC-2002

cDNA, mRNA sequence.

ACCESSION CA829022

VERSION CA829022.1 GI:26457439

KEYWORDS EST.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 620)

REFERENCE

Walbot, V.

Maize ESTs from various cDNA libraries sequenced at Stanford

University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 1114036 row: D column: 07.

FEATURES  
source

Location/Qualifiers  
1. .620  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:4577"  
/clone\_lib="1114 - Unigene IV from Maize Genome Project"  
/note="This library represents the unique genes found in the fourth round of EST sequencing at Stanford University for the maize genome project. Sequences are present from libraries 1091 and 3524. Contigs were assembled using ZmBAssembler and 2 representatives from each contig were selected for the Unigene set. All singlets were also selected."

ORIGIN

Query Match 6.2%; Score 44.4; DB 14; Length 620;  
Best Local Similarity 46.0%; Pred. No. 1.4;  
Matches 150; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 349 TTGGCGCGTTGGCGGTGGCCACTGACACGCGGCGGCTGGAGGTGAGCGTC 408  
|||||  
Db 357 TTGGCGCGTGGCGGAGGTGGAGATGGCGCCCTTGGAGAGCTTGGCGGTATGCTC 298  
|||||  
QY 409 GATAAGCAGCGGTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGGTGAAC 468  
|||||  
Db 297 CTCTGCTAGTAGACGAGGTGGGAGCGGAAGTGGACGTAGCAGGTGGCGGAGGTGATC 238  
|||||  
QY 469 CGGATGCGTATTGACGCGTGTGTTTATCGCGGCGTGGCGCGCAATACGGCGAC 528  
|||||  
Db 237 TCGAGGCGCGCGTGGCGCTCGAAGCTGTAGGAGGCGACGAGGTCCGGGATGAGCCCC 178  
|||||  
QY 529 ACCGAGCGTGGATTTTTCGCGCTGGCGCGTTCGCGGCAAGCTGATCTGGTTCGCGCTG 588  
|||||  
Db 177 TTCGGAGCGCGTACTTGGGAGTAGTGTCTGCGCGCGCTGAGCGAGCTGGCTGAC 118  
|||||  
QY 589 GTGGGTTTCGCGCGACAGCAATGTACGCGCGCTGTCCAGCCCAAGTGTGGCGCTGG 648  
|||||  
Db 117 GCGGTGACCGCTGGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 58  
|||||  
QY 649 ATCAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 674  
|||||  
Db 57 GACATTGCGCGTTGGGATATTGGAT 32  
|||||

RESULT 11

BQ704105/c  
LOCUS 946110C04.y1 946 - tassal primordium prepared by Schmidt lab Zea  
DEFINITION 644 bp mRNA linear EST 16-JUL-2002  
mays cDNA, mRNA sequence.

ACCESSION BQ704105  
VERSION BQ704105.1 GI:21843524  
KEYWORDS EST.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 644)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 946110 row: C column: 04.  
Location/Qualifiers  
1. .644  
/organism="Zea mays"

/mol\_type="mRNA"  
/cultivar="OH43"  
/db\_xref="taxon:4577"  
/tissue\_type="tassels"  
/dev\_stage="just after the transition from vegetative to inflorescence development"  
/lab\_host="XLOLR"  
/clone\_lib="946 - tassal primordium prepared by Schmidt lab"  
/note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI; Site 2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."

ORIGIN

Query Match 6.2%; Score 44.4; DB 13; Length 644;  
Best Local Similarity 46.0%; Pred. No. 1.4;  
Matches 150; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 349 TTGGCGCGTTGGCGGTGGCCACTGACACGCGGCGGCTGGAGGTGAGCGTC 408  
|||||  
Db 364 TTGGCGCTGGAGCGCGGAGAGGTCCGAGATGGCGCCCTTGGAGAGCTTGGCGGTATGCTC 305  
|||||  
QY 409 GATAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGGTGAAC 468  
|||||  
Db 304 CTCTGCTAGTAGACGAGGTGGGAGCGGAAGTGGACGTAGCAGGTGGCGGAGGTGATC 245  
|||||  
QY 469 CGGATGCGTATTGACGCGTGTGTTTATCGCGGCGTGGCGCGCAATACGGCGAC 528  
|||||  
Db 244 TCGAGGCGCGCGTGGCGCTCGTCTGAGAGCTGTAGGAGGCGACGAGTCCGGGATGAGCCCC 185  
|||||  
QY 529 ACCGAGCGTGGATTTTTCGCGCTGGCGCGTTCGCGGCAAGCTGATCTGGTTCGCGCTG 588  
|||||  
Db 184 TTGGCGGAGCGCTACTTGGGAGTAGTGTCTGCGCGCGCTGAGCGAGCTGGCTGAC 125  
|||||  
QY 589 GTGGGTTTCGCGCGACAGCAATGTACGCGCGCTGTCCAGCCCAAGTGTGGCGCTGG 648  
|||||  
Db 124 GCGGTGACCGCTGGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 65  
|||||  
QY 649 ATCAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 674  
|||||  
Db 64 GACATTGCGCGTTGGGATATTGGAT 39  
|||||

RESULT 12

AW927407/c  
LOCUS 945001E08.Y1 945 - Mixed adult tissues from Walbot lab, same as 707  
DEFINITION (SK) Zea mays cDNA, mRNA sequence.

ACCESSION AW927407  
VERSION AW927407.1 GI:8102603  
KEYWORDS EST.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 532)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 945001 row: E column: 08.  
Location/Qualifiers  
1. .532  
source

```

/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db_xref="taxon:4577"
/dev_stage="fully-grown"
/lab_host="DH108"
/clone_lib="945" - Mixed adult tissues from Walbot lab,
same as 707 (SK)
/Note="Organ: tassels, kernal, silk, husk, root, leaf;
Vector: pBAD10; Site 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

```

## ORIGIN

```

Query Match      6.0%; Score 42.8; DB 10; Length 532;
Best Local Similarity 45.7%; Pred. No. 3.4;
Matches 149; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 349 TTGGCGGTTTCGGCGGTTGGGTAAGCCCATGTCACGCGCAGCCGGTGGGGTGGAGGTGAGCGTC 408
    |||||
Db 332 TTGGCGGTTTCGGCGGTTGGGTAAGCCCATGTCACGCGCAGCCGGTGGGGTGGAGGTGAGCGTC 273
    |||||

QY 409 GATAAGCAGCGGTTTGGGTAAGCCCATGTCACGCGCAGCCGGTGGGGTGGAGGTGAGCGTC 468
    |||||
Db 272 CTCTCGTAGTACAGAGTGGGAGCGGAGTGGCGCCCTTGGAGAGCTTCCCGGTTATGGTC 213
    |||||

QY 469 CCGAATGCGTATTTGGACGCGGTTTGTGTTATCGCGCGGCTCGCGCGCAATACGCGCAG 528
    |||||
Db 212 TCGAAGCGCGCGTGGCTCGTGAAGCTGTAGGAGCGGAGCTCGCGGAGTGGCGCC 153
    |||||

QY 529 ACCGAGCGTGAATTTTCGGCGGTTGGCGGTCAGCGCAGCGTATCTGTTCCGCTG 588
    |||||
Db 152 TTGGGAGCGCGTACTTGGGAGTAGTGTCTGCGCGCGGCTGAGCGAGCTGGCTGAC 93
    |||||

QY 589 GTGGGTTTCGGCGCAGCAGATTGTCAAGCGCGCTGCGAGCCCGCAGCGTGTGGCGCTG 648
    |||||
Db 92 GCGGTGACGCTGGCGCGCGGCGCGCGGCGGCGGCGGAGGAGCAGATGCGGGCGCGGA 33
    |||||

QY 649 ATCAACGCTGCTGCGCAGTTGTGAT 674
    |||||
Db 32 GACATTGCGGTTGGGATTTGGGGT 7
    |||||

```

```

RESULT 13
CD861975
LOCUS AZ01.101N14F010125 AZ01 Triticum aestivum cDNA clone AZ01101N14,
DEFINITION mRNA sequence.
ACCESSION CD861975.1 GI:32545791
VERSION CD861975.1
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 538)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.inbio.fr).
FEATURES
source
1. .538
Location/Qualifiers

```

```

/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZ01101N14"
/tissue_type="leaf"
/clone_lib="AZ01"

```

## ORIGIN

```

Query Match      6.0%; Score 42.8; DB 14; Length 538;
Best Local Similarity 47.4%; Pred. No. 3.4; Indels 0; Gaps 0;
Matches 128; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 405 CGTCGATTAAGCAGCGGTTGGGTAAGCCCATGTCACGCGCAGCCGGTGGGGTGGAGGTGAGCGTC 464
    |||||
Db 108 CGTGAGGAGAGCTGTGGCGCTCAAGCCCAAGAGCATGCACTTCGCGCAGCGCGCGG 167
    |||||

QY 465 GAACCCGATGCGTATTTGGAGCGGTTTGTGTTATCGCGCGCTCGCGCGCAATACGG 524
    |||||
Db 168 TCTGCCCTCGCCATCATGAGAGCCGCCCATGAGGGCTCGAGAGAGAGGCTTCTCGCGCGG 227
    |||||

QY 525 CGACACCGGAGCGTGGATTTTCGCGCGCTCGCGCGCAAGCCTGATCTGTTCC 584
    |||||
Db 228 CAATCCATCTTGTCTCGCGCGCGCGCGAGTCGGGACCTTGTATCCAGCTGGC 287
    |||||

QY 585 GCTGCTGGGTTTCGCGCAGCAGCATTTGTCAGCCCGCTGCCAGCCCGCAAGGTGTGGCG 644
    |||||
Db 288 AAGCAAGTTTACGGCGCATCGAAGTGGCGCGCACAGCCAGCAGCCCAAGCTGGAGCT 347
    |||||

QY 645 CTGATCAGCTGCTGCGGCGAGTTGTGAT 674
    |||||
Db 348 CCGAAAAGCTTGGGAGCGGAGCTGCGCAT 377
    |||||

```

## RESULT 14

```

BX425603 885 bp mRNA linear EST 15-MAY-2003
BX425603 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
LOCUS CL0BB0222A07 3-PRIME, mRNA sequence.
ACCESSION BX425603.1 GI:30770486
VERSION BX425603
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 885)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CL0BB0222A07FPL.
FEATURES
source
1. .885
Location/Qualifiers

```

## FEATURES

```

1. .885
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BB0222A07"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
Note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

## ORIGIN

Query Match 5.9%; Score 41.8; DB 13; Length 885;  
Best Local Similarity 0.9%; Pred. No. 7.3;  
Matches 4; Conservative 141; Mismatches 279; Indels 0; Gaps 0;  
75 TGTACTGCTGATTAACACAGGATTAACGCGGAGGACTCATTGCGGTCTCTCTCGTGTG 134  
455 KKK 514  
135 TTTAAATTTCTGACGCTCTTTTGTTCATCGCGGACACCTTGGCGGTGATCTTTGTCCAA 194  
515 KKK 574  
195 TGGCGCGCGCATCGTGTGCGATATATAGCGCTGGGGTGGCATCGCTTACCTGTTATGTTT 254  
575 MNN 634  
255 TGGCGCTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 314  
635 KKK 694  
315 AGAAACAGAACCAACCGTGGCGGATGACACGCTTTTGGCGGTTCGGCGGTGGCCACTGA 374  
695 NNN 754  
375 CACCGCAACCGGGTGGGGTGGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 434  
755 KKK 814  
435 CATGTTGATGCAATCGTGTGCTGACCTGTTGAACCGAATCGTATTTGACGCGTTGT 494  
815 NNN 874  
495 GTTT 498  
875 KKKK 878

RESULT 15  
LOCUS CNS0060N 910 bp DNA linear GSS 03-JUN-1999  
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #  
BACR14J21 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
ACCESSION AL065629  
VERSION AL065629.1 GI:4944698  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 910)  
Genoscope.  
AUTHORS Direct Submission  
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
JOURNAL Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
COMMENT The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazuo Osoegawa and  
Aaron Mammoss in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain y2; cn bw sp, the same strain used for the BDGP's  
P1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).  
Location/Qualifiers

1. .910  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR14J21"  
/clone\_lib="RPCI-98"  
/note="end : T7"  
ORIGIN  
Query Match 5.8%; Score 41.2; DB 29; Length 910;  
Best Local Similarity 14.8%; Pred. No. 11;  
Matches 47; Conservative 150; Mismatches 119; Indels 2; Gaps 1;  
333 GCCCGATGACACGCTTTGGCGGCTTCGCGGTGGCCACGACGCGGCGGTGG 392  
518 GSSSSKGGCGGSGSYGGGGGGYGGTGGGGGGSSSSSSSSSSSSSSSSSSSSSS 577  
393 GGTGAGGTGACGCTGCGATAGCAGCGGGTTTGGTAAAGCCCATGTTGATGCGCAATCGT 452  
578 SSCGCGSCSYSS 637  
453 GCTGACC--TGGTTGAACCCGAATGCGTATTTGGACGCGTGTGTTATCGCGCGGCTC 510  
638 SBTBSSSTBT 697  
511 GCGCGCAATACGCGCACGCGACGCGTGGATTTTCGCGCTGGCGGCTTCGCGCAAGC 570  
698 SSSSSTSCSTKBYSTSTBSYBTBTBTBTBTBTBTBTBTBTBTBTBTBTBTBTBT 757  
571 CTGATCTGTTCCCGCTGGTGGTTCGCGCAGCAGCATGTCACGCCCGCTGTCCACG 630  
758 YTTBBSSTKST 817  
631 CCCAAGGTGTGGCGCTGG 648  
818 SBCTSTSTSSBBSBS 835

Search completed: March 16, 2004, 02:52:15  
Job time : 1958.25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:15:20 ; Search time 3495.31 Seconds  
(without alignments)  
10825.501 Million cell updates/sec

Title: US-09-105-117K-1\_COPY\_1421\_2293

Perfect score: 873

Sequence: 1 gtccataagcagcggttg.....ctgtgagctctgacccgtag 873

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.in.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.ey.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	873	100.0	2374	1	CGLYSEG	X96471 C.glutamici
2	873	100.0	2374	6	AG9393	Sequence 2
3	873	100.0	333150	1	AP005277	AP005277 Coryneb
4	873	100.0	349980	6	AX127147	AX127147 Sequence
5	568	65.1	627	6	AX123538	AX123538 Sequence
6	568	65.1	627	6	BD165655	BD165655 Novel pol
7	568	65.1	1083	6	AX813971	AX813971 Sequence
8	404.2	46.3	993	6	AX067087	AX067087 Sequence
9	329.2	37.7	308650	1	AP005218	AP005218 Coryneb
10	326	37.3	822	6	AX063771	AX063771 Sequence
11	326	37.3	822	6	AX244059	AX244059 Sequence
12	306	35.1	711	6	AX643028	AX643028 Sequence
13	304.4	34.9	712	6	AX643030	AX643030 Sequence
14	303	34.7	708	6	AX123539	AX123539 Sequence
15	303	34.7	708	6	BD165656	BD165656 Novel pol
16	294.6	33.7	349535	1	EX248357	EX248357 Coryneb
17	193.6	22.2	1568	6	E54483	E54483 Heat-resist
18	193.6	22.2	1771	1	AB083133	AB083133 Coryneb
19	141.2	16.2	345783	1	AP003001	AP003001 Mesorhizo
20	136.6	15.6	8472	1	AB008076	AB008076 Agrobacte
21	136.6	15.6	13051	1	AE009111	AE009111 Agrobacte
22	133.6	15.3	302007	1	SC039132	AL939132 Streptomy
23	115	13.2	306250	1	SMES91788	AL591788 Sinorhizo
24	89.6	10.3	298300	1	AP005025	AP005025 Streptomy
25	88.6	10.1	10871	1	AE011965	AE011965 Xanthomon
26	88.4	10.1	303550	1	SC039131	AL939131 Streptomy
27	87.2	10.0	1525	1	PSE1GRA	AE002063 Deinococc
28	86.2	9.9	12257	1	AE002063	AE002063 Escherich
29	84.8	9.7	10336	1	AE005530	AE015317 Shigella
30	83.6	9.6	10278	1	AE015317	AP002563 Escherich
31	83.6	9.6	266658	1	AP002563	AE016988 Shigella
32	83.6	9.6	292906	1	AE016988	AE016766 Escherich
33	82	9.4	303121	1	AE016766	AE000382 Escherich
34	80.4	9.2	11024	1	AE000382	U28377 Escherichia
35	80.4	9.2	141744	1	ECU28377	AC020884 Mus muscu
36	79.4	9.1	110000	2	ACU20884_0	AE016857 Pseudomon
37	78.4	9.0	334108	1	AE016857	AL627274 Salmonell
38	76.4	8.8	256050	1	AL627274	AE016835 Salmonell
39	76.4	8.8	300431	1	AE016835	AE008808 Salmonell
40	74.8	8.6	22929	1	AE008808	AL646059 Ralstonia
41	73.4	8.4	190050	1	AL646059	AE016937 Bacteroid
42	73.4	8.4	305961	1	AE016937	AK012071 Sequence
43	71.6	8.2	5541	6	AR012071	AK025196 Sequence
44	71.6	8.2	5541	6	AR025196	AR038653 Sequence
45	71.6	8.2	5541	6	AR038653	

ALIGNMENTS

RESULT 1

CGLYSEG

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

CGLYSEG 2374 bp DNA linear BCT 19-MAR-2001  
LOCUS C.glutamici lysE and lysG genes.  
DEFINITION X96471  
ACCESSION X96471.1 GI:1729753  
VERSION lysE gene; lysG gene; Lysine export regulator protein; Lysine  
KEYWORDS exporter protein; Lysine governor.  
SOURCE Corynebacterium glutamicum  
ORGANISM Corynebacterium glutamicum  
REFERENCE 1  
AUTHORS Vrljic, M., Sahm, H. and Eggeling, L.  
TITLE A new type of transporter with a new type of cellular function:

JOURNAL L-lysine export from *Corynebacterium glutamicum*  
MEDLINE Mol. Microbiol. 22 (5), 815-826 (1996)  
PUBMED 97126810  
REFERENCE 89711704  
2 (bases 1 to 2374)  
AUTHORS Vrljic, M.M.  
TITLE Direct Submission  
JOURNAL Submitted (07-Mar-1996) M.M. Vrljic, Institut fuer Biotechnologie  
1, Forschungszentrum Juelich, Postfach 1913, D-52425 Juelich, FRG  
FEATURES  
source Location/Qualifiers  
1..2374  
/organism="Corynebacterium glutamicum"  
/mol\_type="genomic DNA"  
/strain="R127"  
/db\_xref="taxon:1718"  
complement (82..954)  
/gene="lysG"  
complement (82..954)  
/gene="lysG"  
/notes="lysine governor"  
/codon\_start=1  
/transl\_table=11  
/product="lysine export regulator protein"  
/protein\_id="CAA65323.1"  
/db\_xref="GI:1729754"  
/db\_xref="GOA:P94632"  
/db\_xref="SWISS-PROT:P94632"  
/translations="MNPIDLTLSIIDEGSFEGLASLALSISPSAVSQRVKALEHVHG  
RVLVSRTQPAKEATGELVQVARKVLLQATKQALSGRLAEPLITIAINADSLSTW  
FFPVNEVAGSGATLTLEDEARTLSLRGDLVIGAVTREAHPVAGGVVVELTWR  
HLAITSPLRDADVMDGLDWAAMPVLRFGKDLVLDRLDGRVDGPGVRRRVSTVS  
ASGFEALRGLGWLPLETQAAPMLKAGEVILLDEIIPDTPMTQWRQVRSRLARL  
TDAVDAALRGLAP"  
1025..1726  
/gene="lysE"  
1025..1726  
/gene="lysE"  
/codon\_start=1  
/transl\_table=11  
/product="lysine exporter protein"  
/protein\_id="CAA65324.2"  
/db\_xref="GI:13397387"  
/db\_xref="GOA:P94633"  
/db\_xref="SWISS-PROT:P94633"  
/translations="MEIFITGLIGASLISIGPNVLVIKQIKRGLIAYLVCLL  
SVDFIFAGTLGVLDLSNAAPIVLDINRWGGIAYLLNFANMAKADANTKVPAPLIE  
ETEPIDDTPLGSSAVADTTRNVRVEVSVDKQVWVPMALVLTMLNPNALDA  
FVFIQVGAQYGDTRGWIIPAGAFASLIWFLVLVFGAALSLRPLSPKVRNWINVVV  
AVVMTALAIKLMLNG"  
complement (1723..2352)  
/gene="orf3"  
complement (1723..2352)  
/gene="orf3"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA65325.1"  
/db\_xref="GI:1729756"  
/db\_xref="SPTFEMBL:P94634"  
/translations="MYALRDIIVASGKALYVIGSSYPGLTAAAEFMWAGCPPLIHQ  
PYSIINRVEEPDGGENLQSAANNGLVIAFSPAQLLTKYLDGIPGSSRASQ  
GKSLSEGMLNINIDVRLKINDIAQERQSLAQMALWYLRQEGYADTVTSALIGA  
SSVEQLDLSLNLNLEFSDAELEAIDRSHDAGINIAKATDSKTREN"  
ORIGIN  
Query Match 100.0%; Score 873; DB 1; Length 2374;  
Best Local Similarity 100.0%; Pred. No. 9.8e-185;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GTCCGATGACGCGGTTGGTAAAGCCCATGTTGATGGCAATCTGCTGACCTGGTTG 60  
Db 1421 GTCCGATGACGCGGTTGGTAAAGCCCATGTTGATGGCAATCTGCTGACCTGGTTG 1480  
Qy 61 AACCCGAATCGCTATTGGACGCGCTTTGTTGTTATCGCGCGCGCTCGCGCGCAATACGCG 120

Db 1481 AACCCGAATCGCTATTGGACGCGCTTTGTTGTTATCGCGCGCGCTCGCGCGCAATACGCG 1540  
Qy 121 GACACCGACGCTGGATTTTCGCCCGCTGCGCGCTGCGCGCAAGCTCTGATCTGTTCCCG 180  
Db 1541 GACACCGACGCTGGATTTTCGCCCGCTGCGCGCTGCGCGCAAGCTCTGATCTGTTCCCG 1600  
Qy 181 CTGCTGGCTTCGCGCGACGACGATTTGTCAGCCCGCTGTCAGCCCGCCCAAGGCTGCGCG 240  
Db 1601 CTGCTGGCTTCGCGCGACGACGATTTGTCAGCCCGCTGTCAGCCCGCCCAAGGCTGCGCG 1660  
Qy 241 TGGATCAACGCTGCTGCGCGCTGTCAGCGCTGTCAGCGCTGTCAGCGCTGTCAGCGCTG 300  
Db 1661 TGGATCAACGCTGCTGCGCGCTGTCAGCGCTGTCAGCGCTGTCAGCGCTGTCAGCG 1720  
Qy 301 GGTAGTTTCGCGCGCTTTGGAAATCGGTGGCTTCGCCCGCAATGTTGATCGCGCGCTG 360  
Db 1721 GGTAGTTTCGCGCGCTTTGGAAATCGGTGGCTTCGCCCGCAATGTTGATCGCGCGCTG 1780  
Qy 361 TGGGAATCTCATCGATCGCTCCAACTCGCGCTCAGAAACTCCAAAGTTGTTGAGTGAA 420  
Db 1781 TGGGAATCTCATCGATCGCTCCAACTCGCGCTCAGAAACTCCAAAGTTGTTGAGTGAA 1840  
Qy 421 TCAAGGCTGTTGTCAGCTGCTCAACTGACGAGCACCAATCAATGCACTGGTCAOGTTA 480  
Db 1841 TCAAGGCTGTTGTCAGCTGCTCAACTGACGAGCACCAATCAATGCACTGGTCAOGTTA 1900  
Qy 481 TCCGCGCGCTACTCTCTTCTGTCGCGACGACCCATGACGAGCGCCATCTGCGGAGTGAC 540  
Db 1901 TCCGCGCGCTACTCTCTTCTGTCGCGACGACCCATGACGAGCGCCATCTGCGGAGTGAC 1960  
Qy 541 TCCGCGCGCTTCTGCGCGATGCTCAATGAGCTTGGCGACCATATCAATATTTGTTCACGTTC 600  
Db 1961 TCCGCGCGCTTCTGCGCGATGCTCAATGAGCTTGGCGACCATATCAATATTTGTTCACGTTC 2020  
Qy 601 AACATGCTCTCAGACGAGGACTTACCTTGGCTGGCGGCGGAAACCTCTGGAATTCGATCG 660  
Db 2021 AACATGCTCTCAGACGAGGACTTACCTTGGCTGGCGGCGGAAACCTCTGGAATTCGATCG 2080  
Qy 661 AGATATTGTCGTCGACGAGCCCTGCGCAAGTGTGAGAAAGCAATGACGCAAGACCA 720  
Db 2081 AGATATTGTCGTCGACGAGCCCTGCGCAAGTGTGAGAAAGCAATGACGCAAGACCA 2140  
Qy 721 TTGTTGGACGCTGATGCAACAGTTCTCACCGTCATCGCGCGCTTCTCCACCCACGA 780  
Db 2141 TTGTTGGACGCTGATGCAACAGTTCTCACCGTCATCGCGCGCTTCTCCACCCACGA 2200  
Qy 781 TTAATGATGATAGCTTGGCTGATGATCAAGAGCGGCGGACCCCTCTCCGCCATGAC 840  
Db 2201 TTAATGATGATAGCTTGGCTGATGATCAAGAGCGGCGGACCCCTCTCCGCCATGAC 2260  
Qy 841 TCAGCGCGCTCGCTGTGAGCTCTGACCGCTAG 873  
Db 2261 TCAGCGCGCTCGCTGTGAGCTCTGACCGCTAG 2293  
RESULT 2  
A93933  
LOCUS Sequence 2 from Patent WO9723597.  
DEFINITION A93933  
ACCESSION A93933  
VERSION A93933.1 GI:6742037  
KEYWORDS  
SOURCE  
ORGANISM  
Corynebacterium glutamicum  
Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
REFERENCE 1 (bases 1 to 2374)  
AUTHORS Vrljic, M. and Eggeling, L.  
TITLE PROCESS FOR THE MICROBIAL PRODUCTION OF AMINO ACIDS BY BOOSTED  
ACTIVITY OF EXPORT CARRIERS  
Patent: WO 9723597-A 2 03-JUL-1997;  
JOURNAL KERNFORSCHUNGSANLAGE JUELICH (DE); VRLIJC MARINA (DE)

FEATURES		Location/Qualifiers	
source	1..2374	/organism="Corynebacterium glutamicum"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:1718"	
ORIGIN			
Query Match 100.0%; Score 873; DB 6; Length 2374;			
Best Local Similarity 100.0%; Pred. No. 9.8e-185;			
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	GTCCGATAGCAGCGGTTGGGTAAAGCCATGTTGATGGCAATCGTGTGACCTGGTTG	60
Db	1421	GTCCGATAGCAGCGGTTGGGTAAAGCCATGTTGATGGCAATCGTGTGACCTGGTTG	1480
QY	61	AACCCGAATCGTATTGGACGCGTTTGTGTTATCGCGCGCTCGCGCGCAATACGCGC	120
Db	1481	AACCCGAATCGTATTGGACGCGTTTGTGTTATCGCGCGCTCGCGCGCAATACGCGC	1540
QY	121	GACACCGGAGCGGTGGATTTTCGCGCTGCGCGCTTGGCGGCAAGCTGATCTGCTCCG	180
Db	1541	GACACCGGAGCGGTGGATTTTCGCGCTGCGCGCTTGGCGGCAAGCTGATCTGCTCCG	1600
QY	181	CTGGTGGGTTTCGCGCGACGACATTTGTCACGCCGCTGTCAGCCCGCAGGTGCGGC	240
Db	1601	CTGGTGGGTTTCGCGCGACGACATTTGTCACGCCGCTGTCAGCCCGCAGGTGCGGC	1660
QY	241	TGGATCAACCTGCTGTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG	300
Db	1661	TGGATCAACCTGCTGTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG	1720
QY	301	GGTTAGTTTTCGCGGGTTTGGATCGGTGGCTTCGCGCCAAATGTTGATGCGCGCTCG	360
Db	1721	GGTTAGTTTTCGCGGGTTTGGATCGGTGGCTTCGCGCCAAATGTTGATGCGCGCTCG	1780
QY	361	TGGGAATCTTCATCGATCGCTCCAACTCGCGCTCAGAAACTCCAAAGTTGTCAGTAA	420
Db	1781	TGGGAATCTTCATCGATCGCTCCAACTCGCGCTCAGAAACTCCAAAGTTGTCAGTAA	1840
QY	421	TCAAGGCTGTGTCAGCTGCTCACTGACGAGCAACCAATCAATGCACTGGTCACGGTA	480
Db	1841	TCAAGGCTGTGTCAGCTGCTCACTGACGAGCAACCAATCAATGCACTGGTCACGGTA	1900
QY	481	TCGCGCGCTGATCTCTCTGCTCGCGACACCAATGCAAGCGCCATCTGCGCAAGTAC	540
Db	1901	TCGCGCGCTGATCTCTCTGCTCGCGACACCAATGCAAGCGCCATCTGCGCAAGTAC	1960
QY	541	TGCGCGCTGCTCTGCGCGCATGTCATGAGCTTGGGACCATATCAATATTTGTCACGTT	600
Db	1961	TGCGCGCTGCTCTGCGCGCATGTCATGAGCTTGGGACCATATCAATATTTGTCACGTT	2020
QY	601	AACATGCTTCAGACAGGAGCTTACCTGCTGCGCGGGAACCTCTCGAATTCATCG	660
Db	2021	AACATGCTTCAGACAGGAGCTTACCTGCTGCGCGGGAACCTCTCGAATTCATCG	2080
QY	661	AGATATTGTCGTCAGCAGCCCTCGCAAGCTGTCGAGCAACATGACCCAGACCA	720
Db	2081	AGATATTGTCGTCAGCAGCCCTCGCAAGCTGTCGAGCAACATGACCCAGACCA	2140
QY	721	TTGTGGCAGCTGACTGCAACAGTTTCTCAACGCTCATCGCGCGGTTCTCCACCCACCA	780
Db	2141	TTGTGGCAGCTGACTGCAACAGTTTCTCAACGCTCATCGCGCGGTTCTCTCCACCCACCA	2200
QY	781	TTAATGATGGAATAGCTTGGCTGATGAATCAGAGCGGGAGCCCTCTCCGCCATGAC	840
Db	2201	TTAATGATGGAATAGCTTGGCTGATGAATCAGAGCGGGAGCCCTCTCCGCCATGAC	2260
QY	841	TCAGCGCTCTCGCTGTGAGCTCTGACCGGTAG	873
Db	2261	TCAGCGCTCTCGCTGTGAGCTCTGACCGGTAG	2293

AP005277/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene





Db 328418 CTGGGGTTTCGGGCGAGCAGCATTTGTCACGCCGCTGTCCAGCCCCAAGGTGTGGCGC 328359  
Qy 241 TGGATCAACCTGCTGCTGGCAGTGTGATGACCGCATGCGCAATCAAACTGATGTGATG 300  
Db 328358 TGGATCAACCTGCTGCTGGCAGTGTGATGACCGCATGCGCAATCAAACTGATGTGATG 328299  
Qy 301 GGTAGTATTTTCGGGGTTCGGAAATCGGTGGCTTCGCCCAAAATGTTGATGCCGCGCTCG 360  
Db 328298 GGTAGTATTTTCGGGGTTCGGAAATCGGTGGCTTCGCCCAAAATGTTGATGCCGCGCTCG 328239  
Qy 361 TGGGAAATCTCATGCAATCGCTCCAACTCGGGGTGAGAAATCTCCAAAGTGTGAGTGAA 420  
Db 328238 TGGGAAATCTCATGCAATCGCTCCAACTCGGGGTGAGAAATCTCCAAAGTGTGAGTGAA 328179  
Qy 421 TCAAGGCTGTTGTCAGCTGCTCAATCTGAGCAAGCAACCAATCAATGCACTGTCACGGTA 480  
Db 328178 TCAAGGCTGTTGTCAGCTGCTCAATCTGAGCAAGCAACCAATCAATGCACTGTCACGGTA 328119  
Qy 481 TCCGCGCGTACTCTCTGCTGCGCAGCAGCCCAATGCAAGCGCAATCTGCGCAAGTGAC 540  
Db 328118 TCCGCGCGTACTCTCTGCTGCGCAGCAGCCCAATGCAAGCGCAATCTGCGCAAGTGAC 328059  
Qy 541 TCCGCGCGTACTCTCTGCTGCGCAGCAGCCCAATGCAAGCGCAATCTGCGCAAGTGAC 600  
Db 328058 TCCGCGCGTACTCTCTGCTGCGCAGCAGCCCAATGCAAGCGCAATCTGCGCAAGTGAC 327999  
Qy 601 AACATGCTCTCAGACAGGAGCTTACCTGCTGGCGGGGAGAGCCCTCTGGAATTCATGCG 660  
Db 327998 AACATGCTCTCAGACAGGAGCTTACCTGCTGGCGGGGAGAGCCCTCTGGAATTCATGCG 327939  
Qy 661 AGATATTGTCCTGAGCAGGAGCTTACCTGCTGGCGGGGAGAGCCCTCTGGAATTCATGCG 720  
Db 327938 AGATATTGTCCTGAGCAGGAGCTTACCTGCTGGCGGGGAGAGCCCTCTGGAATTCATGCG 327879  
Qy 721 TTGTTGGAGCTGACTGCAACAGTCTCAACGCTGATCGCCGCTTCTCCACCCCAAGCA 780  
Db 327878 TTGTTGGAGCTGACTGCAACAGTCTCAACGCTGATCGCCGCTTCTCCACCCCAAGCA 327819  
Qy 781 TTAATGATGGAATAGCTTGGCTGATGAATCAGAGCGGGCAGCCCTCTCCGCGCATGAC 840  
Db 327818 TTAATGATGGAATAGCTTGGCTGATGAATCAGAGCGGGCAGCCCTCTCCGCGCATGAC 327759  
Qy 841 TCAGCGCGCTCCGCTGAGCTCTGGACCGTAG 873  
Db 327758 TCAGCGCGCTCCGCTGAGCTCTGGACCGTAG 327726

RESULT 4  
AX127147/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .349980  
/organism="Corynebacterium glutamicum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1718"  
/note="Seq 1 to long (3,309,400) split in 11, seq 7063  
1.200.001 1.549.980"

AX127147 349980 bp DNA linear PAT 11-MAY-2001  
Sequence 7063 from Patent EP1108790.  
AX127147 AX114121  
AX127147.1 GI:14041135  
Corynebacterium glutamicum  
Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
1  
Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.  
Novel polynucleotides  
Patent: EP 1108790-A 7063 20-JUN-2001;  
KYOWA HAKKO KOGYO CO., LTD. (JP)  
Location/Qualifiers  
1. .349980  
/organism="Corynebacterium glutamicum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1718"  
/note="Seq 1 to long (3,309,400) split in 11, seq 7063  
1.200.001 1.549.980"

ORIGIN

Query Match 100.0%; Score 873; DB 6; Length 349980;  
Best Local Similarity 100.0%; Pred. No. 7.7e-185; Indels 0; Gaps 0;  
Matches 873; Conservative 0; Mismatches 0; Gaps 0;  
Qy 1 GTGATTAAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGGTTG 60  
Db 128548 GTGATTAAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGGTTG 128489  
Qy 61 AACCCGAATGCGTATTGAGACGCGTTTGTGTTATCGGGGGGCTCGGCGGCAATACCGC 120  
Db 128488 AACCCGAATGCGTATTGAGACGCGTTTGTGTTATCGGGGGGCTCGGCGGCAATACCGC 128429  
Qy 121 GACACCGAGCGTGGATTTTTCGCGCTGCGCGCAAGCTGATCTGTTGTTCCCG 180  
Db 128428 GACACCGAGCGTGGATTTTTCGCGCTGCGCGCAAGCTGATCTGTTGTTCCCG 128369  
Qy 181 CTGTGGGTTTCGGCGCAGCAGCATTTGTCAGCCGCTGTGTCAGCCCAAGGTGTGGCGC 240  
Db 128368 CTGTGGGTTTCGGCGCAGCAGCATTTGTCAGCCGCTGTGTCAGCCCAAGGTGTGGCGC 128309  
Qy 241 TGGATCAACGCTGCTGCGGAGTGTGATGACCGCATTTGGCCCACTGAATGTTGATG 300  
Db 128308 TGGATCAACGCTGCTGCGGAGTGTGATGACCGCATTTGGCCCACTGAATGTTGATG 128249  
Qy 301 GGTAGTATTTTCGGGGTTCGGAAATCGGTGGCTTCGCCCAAAATGTTGATGCCGCGCTCG 360  
Db 128248 GGTAGTATTTTCGGGGTTCGGAAATCGGTGGCTTCGCCCAAAATGTTGATGCCGCGCTCG 128189  
Qy 361 TGGGAAATCTCATGCAATCGCTCCAACTCGGGGTGAGAAATCTCAAACTGATGTGATGAA 420  
Db 128188 TGGGAAATCTCATGCAATCGCTCCAACTCGGGGTGAGAAATCTCAAACTGATGTGATGAA 128129  
Qy 421 TCAAGGCTGTTGTCAGCTGCTCAACTGAGCAAGCAACCAATCAATGCACTGTCACGGTA 480  
Db 128128 TCAAGGCTGTTGTCAGCTGCTCAACTGAGCAAGCAACCAATCAATGCACTGTCACGGTA 128069  
Qy 481 TCCGCGCGTACTCTCTGCTGCGCAGCAGCCCAATGCAAGCGCCATCTCGCAAGTGAC 540  
Db 128068 TCCGCGCGTACTCTCTGCTGCGCAGCAGCCCAATGCAAGCGCCATCTCGCAAGTGAC 128009  
Qy 541 TCCGCGCGTACTCTCTGCTGCGCAGCAGCCCAATGCAAGCGCCATCTCGCAAGTGAC 600  
Db 128008 TCCGCGCGTACTCTCTGCTGCGCAGCAGCCCAATGCAAGCGCCATCTCGCAAGTGAC 127949  
Qy 601 AACATGCTCTCAGACAGGAGCTTACCTGCTGGCGGGGAGAGCCCTCTGGAATTCATGCG 660  
Db 127948 AACATGCTCTCAGACAGGAGCTTACCTGCTGGCGGGGAGAGCCCTCTGGAATTCATGCG 127889  
Qy 661 AGATATTGTCCTGAGCAGGAGCTTACCTGCTGGCGGGGAGAGCCCTCTGGAATTCATGCG 720  
Db 127888 AGATATTGTCCTGAGCAGGAGCTTACCTGCTGGCGGGGAGAGCCCTCTGGAATTCATGCG 127829  
Qy 721 TTGTTGGAGCTGACTGCAACAGTCTCAACGCTGATCGCCGCTTCTCCACCCCAAGCA 780  
Db 127828 TTGTTGGAGCTGACTGCAACAGTCTCAACGCTGATCGCCGCTTCTCCACCCCAAGCA 127769  
Qy 781 TTAATGATGGAATAGCTTGGCTGATGAATCAGAGCGGGCAGCCCTCTCCGCGCATGAC 840  
Db 127768 TTAATGATGGAATAGCTTGGCTGATGAATCAGAGCGGGCAGCCCTCTCCGCGCATGAC 127709  
Qy 841 TCAGCGCGCTCCGCTGAGCTCTGGACCGTAG 873  
Db 127708 TCAGCGCGCTCCGCTGAGCTCTGGACCGTAG 127676

RESULT 5  
AX123538/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

AX123538 627 bp DNA linear PAT 11-MAY-2001  
Sequence 3454 from Patent EP1108790.  
AX123538  
AX123538.1 GI:14041026

SOURCE Corynebacterium glutamicum  
ORGANISM Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE  
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.  
TITLE Novel polynucleotides  
JOURNAL Patent: EP 1108790-A 3454 20-JUN-2001;  
KYOWA HAKKO KOGYO CO., LTD. (JP)

FEATURES  
source  
1. .627  
/organism="Corynebacterium glutamicum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1718"

ORIGIN  
Query Match 65.1%; Score 568; DB 6; Length 627;  
Best Local Similarity 100.0%; Pred. No. 1.7e-116;  
Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 GTTTTCGGGGTTTGAATCGGTGGCTTCGCCCAAAATGTTGATGCGGGCGTGGGA 365  
Db 627 GTTTTCGGGGTTTGAATCGGTGGCTTCGCCCAAAATGTTGATGCGGGCGTGGGA 568

QY 366 AATCTCATGATCGCTCCAACTCGGCGTCAGAAAATCCAAAGTTTGTGATGAATCAAG 425  
Db 567 AATCTCATGATCGCTCCAACTCGGCGTCAGAAAATCCAAAGTTTGTGATGAATCAAG 508

QY 426 GCTGTGTCAGTGCCTCAACTGACGAGCACCATCAATGACCTGCTGACGATATCCGC 485  
Db 507 GCTGTGTCAGTGCCTCAACTGACGAGCACCATCAATGACCTGCTGACGATATCCGC 448

QY 486 GCGGTACTCTCTCTGTCGCGAGCACCATCAAGCGCCATCTGCGCAAGTGAATGAT 545  
Db 447 GCGGTACTCTCTCTGTCGCGAGCACCATCAAGCGCCATCTGCGCAAGTGAATGAT 388

QY 546 GCGTTCGGGCGATGATGATGCTGGGAGCACCATCAATGATGATGATGATGATGAT 605  
Db 387 GCGTTCGGGCGATGATGATGCTGGGAGCACCATCAATGATGATGATGATGATGAT 328

QY 606 GCGTTCAGACAGGACTTACCTGCGCGGAGCACCATCAATGATGATGATGATGATGAT 665  
Db 327 GCGTTCAGACAGGACTTACCTGCGCGGAGCACCATCAATGATGATGATGATGATGAT 268

QY 666 TTTGTCGTCAGTGGCTGATGAATCAGAGCGGGCAGCCCTCTCTCGGCATGAATCAG 725  
Db 147 GATGGAATAGTGGCTGATGAATCAGAGCGGGCAGCCCTCTCTCGGCATGAATCAG 88

QY 846 GCGTTCGTCGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873  
Db 87 GCGTTCGTCGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60

RESULT 6  
BD16555/c  
LOCUS BD16555  
DEFINITION Novel polynucleotide.  
ACCESSION BD16555  
VERSION BD16555.1 GI:27871467  
KEYWORDS JP 2002191370-A/3454.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 627)

AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.  
TITLE Novel polynucleotide  
JOURNAL Patent: JP 2002191370-A 3454 09-JUL-2002;  
KYOWA HAKKO KOGYO CO LTD

COMMENT  
OS Corynebacterium glutamicum  
PN JP 2002191370-A/3454  
PD 09-JUL-2002  
PF 15-DEC-2000 JP 2000405096  
PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,  
KEIKO OCHIAI,  
PI HARUHIKO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO  
PI OZAKI  
PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC  
C12N1/15  
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/  
PC 04, C12P13/08,  
PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC  
G01N33/566,  
PC G01N33/569, G01N33/69, G01N37/00//C12P21/08, (C12N1/21, C12R1:15),  
PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15),  
PC C12N15/00,  
PC C12N5/00, C12N15/00  
CC Novel polynucleotide  
FH Key Location/Qualifiers  
FT source 1. .627  
FT /organism="Corynebacterium glutamicum".  
FT Location/Qualifiers  
1. .627  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

ORIGIN  
Query Match 65.1%; Score 568; DB 6; Length 627;  
Best Local Similarity 100.0%; Pred. No. 1.7e-116;  
Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 GTTTTCGGGGTTTGAATCGGTGGCTTCGCCCAAAATGTTGATGCGGGCGTGGGA 365  
Db 627 GTTTTCGGGGTTTGAATCGGTGGCTTCGCCCAAAATGTTGATGCGGGCGTGGGA 568

QY 366 AATCTCATGATCGCTCCAACTCGGCGTCAGAAAATCCAAAGTTTGTGATGAATCAAG 425  
Db 567 AATCTCATGATCGCTCCAACTCGGCGTCAGAAAATCCAAAGTTTGTGATGAATCAAG 508

QY 426 GCTGTGTCAGTGCCTCAACTGACGAGCACCATCAATGACCTGCTGACGATATCCGC 485  
Db 507 GCTGTGTCAGTGCCTCAACTGACGAGCACCATCAATGACCTGCTGACGATATCCGC 448

QY 486 GCGGTACTCTCTCTGTCGCGAGCACCATCAAGCGCCATCTGCGCAAGTGAATGAT 545  
Db 447 GCGGTACTCTCTCTGTCGCGAGCACCATCAAGCGCCATCTGCGCAAGTGAATGAT 388

QY 546 GCGTTCGGGCGATGATGATGCTGGGAGCACCATCAATGATGATGATGATGATGAT 605  
Db 387 GCGTTCGGGCGATGATGATGCTGGGAGCACCATCAATGATGATGATGATGATGAT 328

QY 606 GCGTTCAGACAGGACTTACCTGCGCGGAGCACCATCAATGATGATGATGATGATGAT 665  
Db 327 GCGTTCAGACAGGACTTACCTGCGCGGAGCACCATCAATGATGATGATGATGATGAT 268

QY 666 TTTGTCGTCAGTGGCTGATGAATCAGAGCGGGCAGCCCTCTCTCGGCATGAATCAG 725  
Db 267 TTTGTCGTCAGTGGCTGATGAATCAGAGCGGGCAGCCCTCTCTCGGCATGAATCAG 208

QY 726 GCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785  
Db 207 GCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 148

QY 786 GATGGAATAGTGGCTGATGAATCAGAGCGGGCAGCCCTCTCTCGGCATGAATCAG 845  
Db 147 GATGGAATAGTGGCTGATGAATCAGAGCGGGCAGCCCTCTCTCGGCATGAATCAG 88

QY 846 CGCTCCGCTGTGAGCTCTGACCGTAG 873  
 Db 87 CGCTCCGCTGTGAGCTCTGACCGTAG 60

RESULT 7  
 AX813971/c  
 LOCUS AX813971 1083 bp DNA linear PAT 02-DEC-2003  
 DEFINITION Sequence 211 from Patent WO03040292.  
 ACCESSION AX813971  
 VERSION AX813971.1 GI:38636296  
 KEYWORDS  
 SOURCE Corynebacterium glutamicum  
 ORGANISM Corynebacterium glutamicum  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1  
 AUTHORS Zelder, O., Pompejus, M., Schroeder, H., Kroeger, B., Klopprogge, C. and  
 Haberhauer, G.  
 TITLE Genes encoding for membrane synthesis and membrane transport  
 proteins  
 JOURNAL Patent: WO 03040292-A 211 15-MAY-2003;  
 BASF AKTIENGESSELLSCHAFT (DE)  
 FEATURES  
 LOCATION/Qualifiers  
 1..1083  
 /organism="Corynebacterium glutamicum"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1718"  
 1..>1083  
 /note="unnamed protein product; RXA07013"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAB81915.1"  
 /db\_xref="GI:38636297"

CDS  
 1..1083  
 /translation="MAVVAQPADNRYDDMIYRRVNSGLKLPALISGLWHNFGDKP  
 LSTORSIIHRAFDGVTHTFDLANNYPGAGAFNFRILREDLKSRRDELIISSKAG  
 WMWPGVPGSGSKYLVSLSQSLTRIGLDVYDFYHRRPDPPTPLETMTALRDIV  
 ASGRALYVIGSSGPELTABAEPAEBCPLLIHOPYSILNINWVEEGDDGENILQ  
 SAANGLVIAFPLAQGLITDKYLDIPEGSRSQGSLSSEGLMNNVNDIVRKIND  
 IAQERGSQAOMALAWVLRQEGYADTVTSALIGASSVEQLDNLSDLSNLEFSDAE  
 LEAIDFISHDAGINIWAKATSKTREN"

ORIGIN  
 Query Match 65.1%; Score 568; DB 6; Length 1083;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-116;  
 Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 GTTTTCGCGGGTTTGGAAATCGGTGCGCTTCGCCCAATGTTGATGCGCGGTCTGTGGGA 365  
 Db 1093 GTTTTCGCGGGTTTGGAAATCGGTGCGCTTCGCCCAATGTTGATGCGCGGTCTGTGGGA 1024

QY 366 AATCTCATGATCGCTTCAACTCGCGGTGAGAAACTCAAACTCAAGTTGTTGATGATCAAG 425  
 Db 1023 AATCTCATGATCGCTTCAACTCGCGGTGAGAAACTCAAACTCAAGTTGTTGATGATCAAG 964

QY 426 GCTGTGTCAGTCTCAATGACGAGCACCACCAATCAATGACGCTGCTCAGGTATCCGC 485  
 Db 963 GCTGTGTCAGTCTCAATGACGAGCACCACCAATCAATGACGCTGCTCAGGTATCCGC 904

QY 486 GCCGTACTCTCTCTCGCGAGCACCACCAATGACGAGCACCACCAATGACGAGTGTGCCCC 545  
 Db 903 GCCGTACTCTCTCTCGCGAGCACCACCAATGACGAGCACCACCAATGACGAGTGTGCCCC 844

QY 546 GCGTTCCTGCGGAGTCAATGAGCTTGGGACCAATCAATGATTTGTTCAAGTTCAACAT 605  
 Db 843 GCGTTCCTGCGGAGTCAATGAGCTTGGGACCAATCAATGATTTGTTCAAGTTCAACAT 784

QY 606 GCGTTCAGACAGGACTTACCTGCTGGCGGGAACCTCTGGAATTTCCATCGAGATA 665  
 Db 783 GCGTTCAGACAGGACTTACCTGCTGGCGGGAACCTCTGGAATTTCCATCGAGATA 724

QY 666 TTTGTCCGTGAGCAGGCCCTCGCAAGTGTGTGAGAAAGCAATGACGCCAAGACCATTTGTT 725

Db 723 TTTGTCCGTGAGCAGGCCCTCGCAAGTGTGTGAGAAAGCAATGACGCCAAGACCATTTGTT 664  
 QY 726 GGCAGCTGACTGCAAAAGTTCTCACCGTCATCGCCGCTTCTCCACCAACCAATTAAT 785  
 Db 663 GGCAGCTGACTGCAAAAGTTCTCACCGTCATCGCCGCTTCTCCACCAACCAATTAAT 604

QY 786 GATGATATGCTTGGCTGATGAATCAGAAGCGGCGCAGCCCTCTCCGCGCATGAATCAGC 845  
 Db 603 GATGATATGCTTGGCTGATGAATCAGAAGCGGCGCAGCCCTCTCCGCGCATGAATCAGC 544

QY 846 GCGCTCCGCTGTGAGCTCTGACCGTAG 873  
 Db 543 GCGCTCCGCTGTGAGCTCTGACCGTAG 516

RESULT 8  
 AX067087/c  
 LOCUS AX067087 993 bp DNA linear PAT 24-JAN-2001  
 DEFINITION Sequence 669 from Patent WO0100805.  
 ACCESSION AX067087  
 VERSION AX067087.1 GI:12544795  
 KEYWORDS  
 SOURCE Corynebacterium glutamicum  
 ORGANISM Corynebacterium glutamicum  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1  
 AUTHORS Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Haberhauer, G.  
 TITLE Corynebacterium glutamicum genes encoding proteins involved in  
 membrane synthesis and membrane transport  
 JOURNAL Patent: WO 0100805-A 669 04-JAN-2001;  
 BASF AKTIENGESSELLSCHAFT (DE)  
 FEATURES  
 LOCATION/Qualifiers  
 1..993  
 /organism="Corynebacterium glutamicum"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1718"  
 101..973  
 /note="unnamed protein product; RXN03164"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAC26771.1"  
 /db\_xref="GI:12544796"  
 /translation="MIYEVGNSGLKLPALISGLWHNFGDKPLSTORSIIHRAFDG  
 VTHFDLANNYPGAGAFNFRILREDLKSRRDELIISSKAGWMWPGVPGSGSK  
 YLVSLSQSLTRIGLDVYDFYHRRPDPPTPLETMTALRDIVASGRALYVIGSSG  
 PLETAABAEPAEBCPLLIHOPYSILNINWVEEGDDGENLLQSAANGLVIAFPL  
 AQGLITDKYLDIPEGSRSQGSLSSEGLMNNVNDIVRKINDIAQERGSQAOMALA  
 WVLREQREYVGAGLP"

ORIGIN  
 Query Match 46.3%; Score 404.2; DB 6; Length 993;  
 Best Local Similarity 98.8%; Pred. No. 7e-80;  
 Matches 418; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 453 AGCAACCAATCAATGACGCTGTCACGGTA--TCCGCGCGGTACTCTCTTGTCTCGCGCAGC 510  
 Db 993 AGCAACCAATCAATGACGCTGTCACGGTAATCGGCGCGGTACTCTCTTGTCTCGCGCAGC 934

QY 511 ACCCATCAAGCGCCATCTGCGCAAGTGAATCGCCGCGGTTCCTTGGCGCATGTCATTGAGC 570  
 Db 933 ACCCATCAAGCGCCATCTGCGCAAGTGAATCGCCGCGGTTCCTTGGCGCATGTCATTGAGC 874

QY 571 TTGCGGACCAATATCAATATTTGTCACATGTCACATGTCACATGTCACATGTCACATGTC 630  
 Db 873 TTGCGGACCAATATCAATATTTGTCACATGTCACATGTCACATGTCACATGTCACATGTC 814

QY 631 CTGGCGGGGAAACCTCTGGAATTTCCATCGAGATATTTGTCGCTGAGCAGGCCCTTGCGCA 690  
 Db 813 CTGGCGGGGAAACCTCTGGAATTTCCATCGAGATATTTGTCGCTGAGCAGGCCCTTGCGCA 754

691 AGTGGTGAAGAAGCAATGACGCCAAGACCATTTGTTGGCAGCTGACCTGCAACAGTTCTCA 750  
 753 AGTGGTGAAGAAGCAATGACGCCAAGACCATTTGTTGGCAGCTGACCTGCAACAGTTCTCA 694  
 751 CCGTCATCCGCGGTTCTCCACCAACCAATTAATGATGAATAGCTTGGCTGATGAATC 810  
 693 CCGTCATCCGCGGTTCTCCACCAACCAATTAATGATGAATAGCTTGGCTGATGAATC 634  
 811 AGAAGCGGCGAGCCCTCTCCGCGCATGAATCAGCGCGCTCCCGCTGAGCTCTGGACCG 870  
 633 AGAAGCGGCGAGCCCTCTCCGCGCATGAATCAGCGCGCTCCCGCTGAGCTCTGGACCG 574  
 871 TAG 873  
 573 TAG 571

RESULT 9  
 AP005218/c  
 LOCUS  
 DEFINITION Corynebacterium efficiens YS-314 DNA, complete genome, section 5/11.  
 ACCSSION AP005218 BA000035  
 VERSION AP005218.1 GI:23493016  
 KEYWORDS Corynebacterium efficiens YS-314  
 ORGANISM Corynebacterium efficiens YS-314  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE  
 1 Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Ueda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and Gojobori, T.  
 Comparative Complete Genome Sequence Analysis of the Amino Acid Replacements Responsible for the Thermostability of Corynebacterium efficiens  
 Genome Res. 13 (7), 1572-1579 (2003)

JOURNAL  
 MEDLINE 22723752  
 PUBMED 12840036  
 REFERENCE 2 (bases 1 to 308650)  
 AUTHORS Kawarabayasi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and Director-General of Biotechnology Center.

TITLE  
 JOURNAL Submitted (17-MAY-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan [E-mail: bio@nitech.go.jp, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424] Kawarabayasi, Y. is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, 305-8565 Japan

COMMENT  
 Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan  
 Itoh, T. is at the Japan Biological Information Research Center, Koto-ku, Tokyo, 135-0064 Japan  
 Yamagishi, A. is at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392 Japan  
 Nishio, Y., Ueda, Y. and Sugimoto, S. are at the Ajinomoto Co., Inc., Kawasaki, Kanagawa, 210-8681 Japan  
 The other authors are at the National Institute of Technology and Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.  
 Location/Qualifiers  
 1. 308650  
 /organism="Corynebacterium efficiens YS-314"  
 /mol\_type="genomic DNA"  
 /strain="YS-314"  
 /db\_xref="taxon:196164"  
 complement (83. .241)  
 /note="CE1178"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAC17988.1"  
 /db\_xref="GI:23493017"

translation="MTTGIIILGTGPGDAGLAKIGSDIIPRGVGIHFRFKPPVALQYR  
 TGAGARNQ"  
 365. .742  
 /note="CE1179"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAC17989.1"  
 /db\_xref="GI:23493018"  
 /translation="MEXSQLOSKTQDPAFVDHLSDDAAAFIDNELSRGMHARLH  
 LVHCACRQETARQDITDFVFEKRNENVTAFSDLKAKLMSLANECAPGGAEDAAI  
 QRPESFVAKVSVVRAVRTQGR"  
 774. .1271  
 /note="CE1180, similar to AJ292256-1|CAC21724.1| percent  
 identity: 26 in 155 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="BAC17990.1"  
 /db\_xref="GI:23493019"  
 /translation="MFSNMGVGEIILLFTVGLIVIGPERLPRIMEDLKAATVAARTAI  
 NNAQNLDTPGGERFDETRKPMQAIAQIRSMSPKTAITRALPDGDDSLDSDPDKIM  
 ADTGAQARQVAKTGTTFEHAVERPADRIDPTQSPNTPGKXNDGCKPAGGCVT  
 WTDLI"  
 complement (1353. .2480)  
 /note="CE1181, similar to AL390975-15|CAC01353.1| percent  
 identity: 63 in 370 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative ATP-binding protein Mrp"  
 /protein\_id="BAC17991.1"  
 /db\_xref="GI:23493020"  
 /translation="MMSQVTSVRSALSRVSDPEIGKPTTEIGMVKVSAIDGSDVQV  
 EVMITTAACPMKTTIVSTAAIKDIEGVGVQVTTVDMSDEQRRLRSLRGNTADP  
 VIPAQSGTTRVAVASGKGVKSMVTNLAALAAKGLSVGLIDADITYGHSVPFG  
 LGSAPRHQVDDIMPQAHGVKMSIAHFVEGNAPVVMRGPMLRHAIQOPLSDVFWG  
 DLDLIDLPFGTGIIATIAQLLPAEALLIVTTTPQAARADVAERAGSISLTQKQVA  
 GVIENNSAMPDGTMEVFGSGGKIAQRLSMLTGEDVPVLGVSPLDPSLRVGGDV  
 GNPIASEPYSPTAAINATADTLAKRKVSIVGKTLGLGVK"  
 complement (2513. .3271)  
 /note="CE1182"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAC17992.1"  
 /db\_xref="GI:23493021"  
 /translation="MTTAPWADAFHVGWDLPOILEVDHTEVAHLOEYFATDPRT  
 CPMVTGSLNRLGGGDAQPVNRNITDLSLLDVPVGLQVQLLEPPEDAQSI  
 DLEWVRDYTAELPYDLQIIGILASIPTDVGLASKAAEHLDAAGNLLMTALRRNG  
 FGPRVSKLIARRRPLPFIDSAIKRQIINTSVGVFENFHTVLAEDKALSRHLKGI  
 RTEADQSGNSALADLTIRFVQVVMADHKRT"  
 complement (3282. .3836)  
 /note="CE1183, similar to AE007003-3|AAK45527.1| percent  
 identity: 57 in 159 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="BAC17993.1"  
 /db\_xref="GI:23493022"  
 /translation="MADYSRDLDTPLVGTGRRFRKDDDTGVAEAKVAFRFGTQGY  
 LPWQTFIVIIWMINIGFANWDPYFILLNIAFSTQAAVAAPLILLAAQNRQEDRK  
 STILADRRARETKATBELRELAVGLAVGEMVTRDYLRLHLLDGLRFLERLEAKL  
 DDVAASNDPDLNAPDTQGVVD"  
 complement (3829. .5124)  
 /note="CE1184, similar to AE007003-4|AAK45528.1| percent  
 identity: 52 in 417 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="BAC17994.1"  
 /db\_xref="GI:23493023"  
 /translation="MRGITRLYAGRLAGWVRGPDITVIGRVVVVTTIRSGQDSRV  
 LGLVVELVNNKRIFLEPLMLRIATAEPKEVNILSGSVLSRSPKARTGELQVIGITGK

VHTDPELEQHGCRVAELADVELSEQTRTRDWISRVALLGERPKFGRRTYIAPWHS  
VHGITAGAGETQTAELIAGFEDMRAADVAKELYGLPKPQDVVAGELDDRLADIL  
QELSDEQALIELTIDIERAADVLEEMDDADLLGLGEDPDRAGVLLBELMDPEESAP  
VRMGFNPTVGTALMTPEPLTMDPSTVAEALAMARNPDLPTSASLVFVVRPEPTAT  
PTGKVLGCVLQKLLREPPSLVGVILDPDLPPLVANDSQETAARFFATYVNLVCGPVL  
DENHGLICAVAVDDLDHMLPEDMDAGIRPGVKEETLG"  
5264. 5782  
/note="CELL185, similar to AB007003-6|AAK45530.1| percent  
identity: 60 in 155 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17995.1"  
/db\_xref="GI:23493024"  
/translation="MSMQATSGGNNRAELGLKKRPEGVGSPQTVBEAQRADVLMS  
DNEFPAELTITVGLVMEVETVGRLTWGRVLGAGASAWLGLFGLVGMVLSNLF  
SLLTGLVMGVVFGVVAAPVWASRGKRDFTSATQIVAGRYDILCRDRAEARDMI  
ARNMGLCHPRQ"  
5805. 6500  
/note="CELL186"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17996.1"  
/db\_xref="GI:23493025"  
/translation="MRTTGVKALIGLVAVLSSLLVAGGSVDNRPPEPRQVAREQX  
ENLYATDVLGNVISELYTILGNRGPVIRHEDNLTADFPDLNEGSDLVIG  
CTGALLEVNPVLAGEEYLAANVAGETDINSGRDRTFDALIGLSPRLDADP  
SNAQCPGANGLPQNIPIYRNPTFSRADREILLNWSAGAITTSLELBEAKELGRS  
RELVEFVASKGG"  
6660. 7373  
/note="CELL187, similar to AB007262-3|AAK65408.1| percent  
identity: 32 in 235 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC17997.1"  
/db\_xref="GI:23493026"  
/translation="METILIGVLLFVACVCGALGFLGMLAAPILAMPDILPGLLI  
LLAFGLSLATWADRGVEVWPVGMVSLGKIPGVSLGTWVAVMPVAGLKVLATAV  
ILGTLISLVSPGHRNSFLAGASGGLGTTTATGGPPLALNRSMSPERVGRGLS  
VCFVIGSISALTGTAGALGWHMLQAALVYAPAVIAGVFLSGVNVKHLNRLLIFLS  
VTLISLGSWITQAGIF"  
7592. 8149  
/note="CELL188"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17998.1"  
/db\_xref="GI:23493027"  
/translation="MGRGDGFPAERFPADPRCKRTIKMQWTKGATAAAALLLV  
GGGTFMSNDEVTSQDAIAGNALSAAGKWTNLNGQDVDRNVNAKVPDGLTV  
FTQGISYETSGNLSATLTTPGSTATADAALATLNDITITIDJGRLPSTITG  
SGNTYFEGVSGSVALSISQAA"  
complement (8190. 9443)  
/note="CELL189, similar to AF024666-23|AAG03374.1| percent  
identity: 84 in 408 aa"  
/codon\_start=1  
/transl\_table=11  
/product="transposase"  
/protein\_id="BAC17999.1"  
/db\_xref="GI:23493028"  
/translation="WAGPYSDPTTLDLLAQAQSPDLRMWLMQGFNOILSTQADQ  
ALTTVATCYLKGVSRRMDLVATLGINNLSKQSDMAKODQDVEEFKPLDTGT  
PLYVSDALTMKRGGRVYKTVLLATGVNAEGYRELLGQVATRESVNSVGFPR  
DLKAGNEVLVTSADALGHTQIAEVLNPNASQRCRTHFSKNLYGMVSKQWPTLS  
AMPHETIFQPDQAQSVMGAREVAVAFCEQKFPDVADYLEALDELLAFATNAPKSVTKV  
WSNPTERANREIRRTVGIPEPDAVRLVGLNVAEQHDDWIQKRYNSLTSLAQ  
TRAWNTAIIIDSDITSIIRKQVA"  
complement (11439. 15143)  
/gene="odha"  
complement (11439. 15143)

Query Match 37.7%; Score 329.2; DB 1; Length 308650;  
Best Local Similarity 74.4%; Pred. No. 3.2e-63;  
Matches 415; Conservative 0; Mismatches 143; Indels 0; Gaps 0;  
QY 316 GTTITGGAATCGGTGGCCCTTCGCCCAAAATGTTGATGCGCGCGTCTGTGGAAATCTCATTCG 375  
DB 187204 GTCTTGAATTCGGTGGCCCTTCGCCCAAGATGTTTACCCCGCGTCATGGCGCATCTCATTCG 187145  
QY 376 ATCCGCTCCAACTCCGGCGTCAGAAAACCTCAAGTTTGTGAGTGAATCAACAGGCTGTTCCTCC 435  
DB 187144 ATGGCCCTCCAGTTCCGCTCGTGAATTCAGGTTGGCCAGGGCGTCCGAGGTGGCATTCG 187085  
QY 436 AGCTGCTCAACTGACGAAGACCAATCAATGCACTGGTCACTCCGCGCGCTACTCTCT 495  
DB 187084 AGCTGCTCAAGCTCGAGGACCCGATGAGCGCGAGGTGACCGTCTCTCCACGAGGAG 187025  
QY 496 CTTTGTCTCGCGACGACCCATGCAAGCGCCATCTCGCAAGTGACTGCGCGCGTTCCTGG 555  
DB 187024 CCTTGTTCGCGAGGACCCAGCGCATCTGCGCGAGGACTGACCGCGCTCTCTGA 186965  
QY 556 GCGATGCTATGAGCTTGGCGACCATATCAATATTTTACGTTTCAACATGCCCTCAGAC 615  
DB 186964 GCGATGCTTGAAGCGCGCCATCCGATGCTCTCTCGACAGCATGCCCTCAGAC 186905  
QY 616 AGGCACTTACCTGCTGCGCGCGGAAACCTCTTGGAAATTCATCGAGATATTTTTCGCTG 675  
DB 186904 AGCGACTTGGCTTGGTGGCAGCGGACCCCTCGGGGACACCGTTCGAGTAGCGTTCGCTG 186845  
QY 676 AGCAGGCGCTCGCAAGTGGTGAGAAAGCAATGACCAAGACCATTTGTGGCACTGAC 735  
DB 186844 AGCAGGCGCTCGCCAAACGCGGAGAGCGGATGACCCCGTTCGCTGTCGCGCGCC 186785  
QY 736 TCGAACAGTTCTTACCGCTTCGCGCGGTTCTCCACCAACGATTATATGATGAATAG 795  
DB 186784 TGAGCAGCGACTCCCTCTATCACCCGTTCTCTCCACCGTTCTCCACCGTTGACAGGAGTAG 186725  
QY 796 CTTGCTGATGAATCAGAAAGCGCGCGCCCTCTCTCGGCATGAATCAAGCGCGCTCCGCT 855  
DB 186724 CTCGCTGGTGGATGAGCAGCGCGGACCCCTCTCTCGGCATGAATCCGCGCGCTCGCG 186665  
QY 856 GTGAGCTCTGACCGTAG 873  
DB 186664 GTGAGATCCGACCGTAG 186647  
RESULT 10  
AX063771 822 bp DNA linear PAT 24-JAN-2001  
LOCUS Sequence 53 from Patent WO0100843.  
DEFINITION AX063771  
ACCESSION AX063771  
VERSION AX063771.1 GI:12541483  
KEYWORDS Corynebacterium glutamicum  
SOURCE Corynebacterium glutamicum  
ORGANISM Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
REFERENCE 1  
AUTHORS Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Habermeyer, G.  
TITLE corynebacterium glutamicum genes encoding metabolic pathway  
JOURNAL proteins  
PATENT: WO 0100843-A 53 04-JAN-2001;  
BASF AKTIENGESSELLSCHAFT (DE)  
FEATURES  
source  
1. 822  
/organism="Corynebacterium glutamicum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1718"  
101. 802  
/note="unnamed protein product; RXA01394"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC25127.1"  
CDS

AVVMTALAIKLMMLMG"

\_\_\_\_\_



```

Db 406 GTCGATAGCAGCGGGTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTG 465
Qy 61 AACCCGAATGCGTATTGAGACGGGTTTGTGTTTATCGGCGGGCGTGGGCGCGCAATACGGC 120
Db 466 AACCCGAATGCGTATTGAGACGGGTTTGTGTTTATCGGCGGGCGTGGGCGCGCAATACGGC 525
Qy 121 GACACCGAGCGGTGGAATTTGCGCGCTGGCGGTTTCGGCGAGCCTGATCTGGTTCCCG 180
Db 526 GACACCGAGCGGTGGAATTTGCGCGCTGGCGGTTTCGGCGAGCCTGATCTGGTTCCCG 585
Qy 181 CTGCTGGGTTTCGGCGCAGCAGCATTTGTCAGCCCGCTGTCAGCCCGCAAGCTGTGGCGC 240
Db 586 CTGCTGGGTTTCGGCGCAGCAGCATTTGTCAGCCCGCTGTCAGCCCGCAAGCTGTGGCGC 645
Qy 241 TGGATCAACGTCGTGTCGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATTTGATG 300
Db 646 TGGATCAACGTCGTGTCGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATTTGATG 705
Qy 301 GGTAG 306
Db 706 GGTAG 711

RESULT 13
AX643030
LOCUS AX643030 712 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 9 from Patent EP1266986.
ACCESSION AX643030
VERSION AX643030.1 GI:28550160
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
1
AUTHORS Gunji, Y. and Yasueda, H.
TITLE Method for producing L-lysine or L-arginine by using methanol
JOURNAL assimilating bacterium
PATENT: EP 1266986-A 9 18-DEC-2002;
AJINOMOTO CO., INC. (JP)
FEATURES
source
1. .712
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
CDS
1. .375
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD67812.1"
/db_xref="GI:28550161"
/translation="MWIMEIFITGLLGLGALLSLISGPNVLVKGIKRSGLIANLVLY
CLISDFVFIATGTLGVLLSNAPIVLDIMRWGGIAYLLFWFAWAAKDMTNKVEAPQ
IIBETEPTVPDPTPLGVFGGH"

ORIGIN
Query Match 34.9%; Score 304.4; DB 6; Length 712;
Best Local Similarity 99.7%; Pred. No. 1.5e-57;
Matches 305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTCGATAGCAGCGGGTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTG 60
Db 407 GTCGATAGCAGCGGGTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTG 466
Qy 61 AACCCGAATGCGTATTGAGACGGGTTTGTGTTTATCGGCGGGCGTGGGCGCGCAATACGGC 120
Db 467 AACCCGAATGCGTATTGAGACGGGTTTGTGTTTATCGGCGGGCGTGGGCGCGCAATACGGC 526
Qy 121 GACACCGAGCGGTGGAATTTGCGCGCTGGCGGTTTCGGCGAGCCTGATCTGGTTCCCG 180
Db 527 GACACCGAGCGGTGGAATTTGCGCGCTGGCGGTTTCGGCGAGCCTGATCTGGTTCCCG 586

```

```

Qy 181 CTGCTGGGTTTCGGCGCAGCAGCATTTGTCAGCCCGCTGTCAGCCCGCAAGCTGTGGCGC 240
Db 587 CTGCTGGGTTTCGGCGCAGCAGCATTTGTCAGCCCGCTGTCAGCCCGCAAGCTGTGGCGC 646
Qy 241 TGGATCAACGTCGTGTCGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATTTGATG 300
Db 647 TGGATCAACGTCGTGTCGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATTTGATG 706
Qy 301 GGTAG 306
Db 707 GGTAG 712

RESULT 14
AX123539
LOCUS AX123539 708 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 3455 from Patent EP1108790.
ACCESSION AX123539
VERSION AX123539.1 GI:14041027
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
1
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 3455 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
source
1. .708
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

ORIGIN
Query Match 34.7%; Score 303; DB 6; Length 708;
Best Local Similarity 100.0%; Pred. No. 3.1e-57;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGATAGCAGCGGGTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTG 60
Db 406 GTCGATAGCAGCGGGTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTG 465
Qy 61 AACCCGAATGCGTATTGAGACGGGTTTGTGTTTATCGGCGGGCGTGGGCGCGCAATACGGC 120
Db 466 AACCCGAATGCGTATTGAGACGGGTTTGTGTTTATCGGCGGGCGTGGGCGCGCAATACGGC 525
Qy 121 GACACCGAGCGGTGGAATTTGCGCGCTGGCGGCTTTCGGCGAGCCTGATCTGGTTCCCG 180
Db 526 GACACCGAGCGGTGGAATTTGCGCGCTGGCGGCTTTCGGCGAGCCTGATCTGGTTCCCG 585
Qy 181 CTGCTGGGTTTCGGCGCAGCAGCATTTGTCAGCCCGCTGTCAGCCCGCAAGCTGTGGCGC 240
Db 586 CTGCTGGGTTTCGGCGCAGCAGCATTTGTCAGCCCGCTGTCAGCCCGCAAGCTGTGGCGC 645
Qy 241 TGGATCAACGTCGTGTCGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATTTGATG 300
Db 646 TGGATCAACGTCGTGTCGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATTTGATG 705
Qy 301 GGT 303
Db 706 GGT 708

RESULT 15
BD165656
LOCUS BD165656 708 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel polynucleotide.
ACCESSION BD165656
VERSION BD165656.1 GI:27871468
KEYWORDS JP 2002191370-A/3455.

```

SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 708)  
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayaashi,M., Ochiai,K.,  
Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.  
TITLE Novel polynucleotide  
JOURNAL Patent: JP 2002191370-A 3455 09-JUL-2002;  
KICWA HAKO KOGYO CO LTD  
OS Corynebacterium glutamicum  
PN JP 2002191370-A/3455  
PD 09-JUL-2002  
PF 15-DEC-2000 JP 2000405096  
PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,  
PI KEIKO OCHIAI,  
PI HARUHIKO YOKOI,NAKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO  
PI OZAKI  
PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC  
C12N1/15,  
PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/04,  
C12P13/08,  
PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53,PC  
G01N33/566,  
PC G01N33/569,G01N33/68,G01N37/00//C12P21/08,(C12N1/21,C12R1:15),  
PC (C12N1/21,C12R1:13),(C12N1/21,C12R1:01),(C12P13/08,C12R1:15),  
PC C12N15/00,  
PC C12N5/00,C12N15/00  
CC Novel polynucleotide  
FH Key Location/Qualifiers  
FT source 1..708  
FT /organism='Corynebacterium glutamicum'.  
Location/Qualifiers  
1..708  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

ORIGIN

Query Match 34.7%; Score 303; DB 6; Length 708;  
Best Local Similarity 100.0%; Pred. No. 3.1e-57;  
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGTTG 60  
|||  
Db 406 GTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGTTG 465  
|||  
Qy 61 AACCCGATGCGTATTGGAGCGTTTGTTTATCGCGCGGTGGCGCCCATACGGC 120  
|||  
Db 466 AACCCGATGCGTATTGGAGCGTTTGTTTATCGCGCGGTGGCGCCCATACGGC 525  
|||  
Qy 121 GACACGGAGCGGTGATTTTCGCGGTGGCGGTTCGCGCAAGCTGATCTGGTCCCG 180  
|||  
Db 526 GACACGGAGCGGTGATTTTCGCGGTGGCGGTTCGCGCAAGCTGATCTGGTCCCG 585  
|||  
Qy 181 CTGGTGGTTTCGGCGCAGCAGCATTTGCAAGCGGTGTCAGCCCAAGGTGGCGC 240  
|||  
Db 586 CTGGTGGTTTCGGCGCAGCAGCATTTGCAAGCGGTGTCAGCCCAAGGTGGCGC 645  
|||  
Qy 241 TGGATCAAGCTGCTGGCAGTTGATGACCGATGGCCATCAAACTGATGTCATG 300  
|||  
Db 646 TGGATCAAGCTGCTGGCAGTTGATGACCGATGGCCATCAAACTGATGTCATG 705  
|||  
Qy 301 GGT 303  
|||  
Db 706 GGT 708  
|||

Search completed: March 15, 2004, 22:02:04  
Job time : 3498.31 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	873	100.0	2374	2	AAT96816	DNA encod
2	873	100.0	2374	9	ADB66196	DNA fragm
3	873	100.0	34980	5	AAH68528	C glutami
C	568	65.1	527	5	AAB68419	C glutami
C	404.2	46.3	993	4	AAF68077	Corynebact
6	326	37.3	822	4	AAF71779	Corynebact
7	326	37.3	822	4	AAAS96098	C. glutam
8	306	35.1	711	8	ACC80941	Lyse prot
9	304.4	34.9	712	8	ACC80942	LyseE24 pr
10	303	34.7	708	5	AAH68420	C glutami
C	253.6	29.0	1095	7	ACA29651	C glutami
12	193.6	22.2	1568	4	AAH45375	C. thermo
C	105.8	12.1	1041	7	ACA25567	Prokaryot
C	96.8	11.1	1107	7	ACA26520	Prokaryot
C	96.2	11.0	1041	7	ACA24017	Prokaryot
C	80.4	9.2	1038	7	ACA32095	Prokaryot
C	79.4	9.1	1041	7	ACA19150	Prokaryot
C	78.2	9.0	1035	7	ACA45509	Prokaryot
C	76.4	8.8	999	7	ACA51539	Prokaryot
C	74.8	8.6	999	7	ACA50894	Prokaryot
20	71.6	8.2	5541	2	AAQ55755	Escherich
C	69.2	7.9	349980	6	ABQ81844	Abq81844
C	65.4	7.5	9988	6	ABK2786	Bifidobac
C	65.4	7.5	9988	6	ABK2786	Bifidobac

XX CC This DNA, isolated from *Corynebacterium glutamicum*, contains the *lysE*,  
CC *lysE* and *orf3* genes. *lysE* and *lysE* encode a lysine transport regulatory  
CC protein and an export protein, respectively. Microbial production of  
CC amino acids (A) is improved by increasing the export-carrier activity  
CC and/or the export gene expression in a microorganism that produces (A).  
CC The method is specifically used to increase production of lysine, used as  
CC an animal feed additive. Other (A) are variously useful as  
CC pharmaceuticals, condiments and intermediates for fine chemicals. This  
CC method increases the amount of (A) secreted into the culture medium.  
CC Export of (A) has been found to depend on a single gene. NB. This  
CC sequence has been created from the information given in table 2 of the  
CC specification  
XX SQ Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 873; DB 2; Length 2374;  
Best Local Similarity 100.0%; Pred. No. 4.9e-246;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GTCGATAAGCAGCGGTTGGGTAAAGCCATGTTGATGGCAATCGTGCTGACCTGGTTG 60  
DB 1421 GTCGATAAGCAGCGGTTGGGTAAAGCCATGTTGATGGCAATCGTGCTGACCTGGTTG 1480  
  
QY 61 AACCCGATGCGTATTGACGCGTTGTTGTTATCGCGCGCTGCGCGCAATACGCG 120  
DB 1481 AACCCGATGCGTATTGACGCGTTGTTGTTATCGCGCGCTGCGCGCAATACGCG 1540  
  
QY 121 GACACCGACGCGTGGATTTTCGCCGCTGGCGCGTTTCGGCGAAGCCTGATCTGGTTCCG 180  
DB 1541 GACACCGACGCGTGGATTTTCGCCGCTGGCGCGTTTCGGCGAAGCCTGATCTGGTTCCG 1600  
  
QY 181 CTGGTGGTTTCGGCGCAGCAGCATTTGTCAGCGCCGCTGTCAGCGCCCAAGTGTGGCG 240  
DB 1601 CTGGTGGTTTCGGCGCAGCAGCATTTGTCAGCGCCGCTGTCAGCGCCCAAGTGTGGCG 1660  
  
QY 241 TGGATCAAGTTCGTCGCGAGTTCGATGACCGCATTCGCCCATCAACTGATGTTGATG 300  
DB 1661 TGGATCAAGTTCGTCGCGAGTTCGATGACCGCATTCGCCCATCAACTGATGTTGATG 1720  
  
QY 301 GGTAGTTTCGCGCGTTTGGAAATCGGTGGCTTCGCCCAATGTTGATGCGCGCTG 360  
DB 1721 GGTAGTTTCGCGCGTTTGGAAATCGGTGGCTTCGCCCAATGTTGATGCGCGCTG 1780  
  
QY 361 TGGAAATCTCATGATCGCTTCAACTCGCGCTGAGAACTCCAGTGTGATGAA 420  
DB 1781 TGGAAATCTCATGATCGCTTCAACTCGCGCTGAGAACTCCAGTGTGATGAA 1840  
  
QY 421 TCAAGCTGTTGTCAGCTGCTCACTGACGAGCAACCAATCAATGCACTGTCAGGTA 480  
DB 1841 TCAAGCTGTTGTCAGCTGCTCACTGACGAGCAACCAATCAATGCACTGTCAGGTA 1900  
  
QY 481 TCGCGCGCTACTCTCTCTCGCGCAGCACCATGCAAGCGCCATCTGCGCAAGTGAC 540  
DB 1901 TCGCGCGCTACTCTCTCTCGCGCAGCACCATGCAAGCGCCATCTGCGCAAGTGAC 1960  
  
QY 541 TGCCCGGTTCTCGCGCATGTCATTGAGCTTGGACCATATCAATATTGTCAGTTTC 600  
DB 1961 TGCCCGGTTCTCGCGCATGTCATTGAGCTTGGACCATATCAATATTGTCAGTTTC 2020  
  
QY 601 AACATGCTCTGACAGCGGACTTACCTGCTGGCGGGAACCTCTGGAATTCATCG 660  
DB 2021 AACATGCTCTGACAGCGGACTTACCTGCTGGCGGGAACCTCTGGAATTCATCG 2080  
  
QY 661 AGATATTGTCGTCAGCAGCGCTCTCGCAAGTGTGAGAAAGCAATGACGCCAAGCA 720  
DB 2081 AGATATTGTCGTCAGCAGCGCTCTCGCAAGTGTGAGAAAGCAATGACGCCAAGCA 2140  
  
QY 721 TTGTTGCGAGCTGATGCAACAGATTCTCACCGTATCGCGGTTCTCCACCAACGA 780  
DB 2141 TTGTTGCGAGCTGATGCAACAGATTCTCACCGTATCGCGGTTCTCCACCAACGA 2200  
  
QY 781 TTAATGATGGAATAGTTGGCTGATGAATCAGAAGCGGCGGCTCTCTCGCGCATGAAC 840

DB 2201 TTAATGATGGAATAGTTGGCTGATGAATCAGAAGCGGCGGCTCTCTCGCGCATGAAC 2260  
QY 841 TCAGCCGCTTCGCTGCTGAGCTCTGAGCTCTGAGCCGTAG 873  
DB 2261 TCAGCCGCTTCGCTGCTGAGCTCTGAGCCGTAG 2293  
  
RESULT 2  
ADB66196  
ID ADB66196 standard; DNA; 2374 BP.  
XX AC ADB66196;  
XX DT 04-DEC-2003 (first entry)  
XX DE DNA fragment containing C. glutamicum *lysG* and *lysE* genes.  
XX KW L-arginine production; coryneform bacteria; *lysE*; arginine repressor;  
XX KW argR; liver function promoting agent; amino acid infusion;  
XX KW amino acid pharmaceutical; *lysG*; ds.  
XX OS *Corynebacterium glutamicum*.  
XX FH Key Location/Qualifiers  
XX CDS 1025..1726  
XX FT /\*tag= a  
XX FT /product= "Protein encoded by *lysE* gene"  
  
XX US2003113899-A1.  
XX PN 19-JUN-2003.  
XX PP 17-JUL-2002; 2002US-00196232.  
XX PR 25-JUL-2001; 2001JP-00224586.  
XX PA (AJIN) AJINOMOTO CO INC.  
XX PI Yamaguchi M, Ito H, Gunji Y, Yasueda H;  
XX WIPI; 2003-708853/67.  
XX P-PSDB; ADB66197.  
  
A microorganism comprising enhanced expression of the *lysE* gene is useful  
for enhanced production of L-arginine.  
  
Example 4; Page 26-27; 36pp; English.  
  
The present invention relates to a method for producing L-arginine in a  
microorganism (e.g. coryneform bacteria) that has L-arginine producing  
ability and has been modified for enhanced expression of the *lysE* gene.  
The microorganism is also modified so that an arginine repressor (*argR*)  
does not function normally. The method of the invention is useful for the  
enhanced production of L-arginine which is useful in liver function  
promoting agents, amino acid infusion and comprehensive amino acid  
pharmaceuticals. The present sequence represents a DNA fragment  
containing *Corynebacterium glutamicum lysG* and *lysE* genes. Note: The  
present sequence is given as SEQ ID No:24 in the Sequence Listing but is  
referred to as SEQ ID No:25 in the rest of the specification.  
  
Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 873; DB 9; Length 2374;  
Best Local Similarity 100.0%; Pred. No. 4.9e-246;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GTCGATAAGCAGCGGTTGGGTAAAGCCATGTTGATGGCAATCGTGCTGACCTGGTTG 60  
DB 1421 GTCGATAAGCAGCGGTTGGGTAAAGCCATGTTGATGGCAATCGTGCTGACCTGGTTG 1480  
  
QY 61 AACCCGATGCGTATTGACGCGTTGTTGTTATCGCGCGCTGCGCGCAATACGCG 120

Db 1481 AACCCGATGCGTATTTGGACGGCTTTGTTTATCGCGGGCTGCGCGCAATACGCG 1540  
Qy 121 GACACCGGACGCTGATTTTCGCGCTGCGCGCTTCCGCGCAAGCTCTGATCTGTTCCCG 180  
Db 1541 GACACCGGACGCTGATTTTCGCGCTGCGCGCTTCCGCGCAAGCTCTGATCTGTTCCCG 1600  
Qy 181 CTGCTGGGTTTCGCGCGACGAGCATTTGTCACGCCGCTGTCACGCCCAAGCTGTGCGCG 240  
Db 1601 CTGCTGGGTTTCGCGCGACGAGCATTTGTCACGCCGCTGTCACGCCCAAGCTGTGCGCG 1660  
Qy 241 TGGATCAACGCTGCTGCTGGCAGTTGTGATGACCGCATGCGCCATCAAACTGATGTGATG 300  
Db 1661 TGGATCAACGCTGCTGCTGGCAGTTGTGATGACCGCATGCGCCATCAAACTGATGTGATG 1720  
Qy 301 GGTAGTTTTCGCGGGTTTGGATCGGTGCGCTTCCGCCAATGTTGATCGCGGCTG 360  
Db 1721 GGTAGTTTTCGCGGGTTTGGATCGGTGCGCTTCCGCCAATGTTGATCGCGGCTG 1780  
Qy 361 TGGGAATCTCATCGCTCCCACTCGCGCTCAGAAATCTCCAAAGTTGTGAGTGAA 420  
Db 1781 TGGGAATCTCATCGCTCCCACTCGCGCTCAGAAATCTCCAAAGTTGTGAGTGAA 1840  
Qy 421 TCAAGGCTGTTCAGCTGCTCAATGACGAGCAATCAATGACATGCTGTCAGGTA 480  
Db 1841 TCAAGGCTGTTCAGCTGCTCAATGACGAGCAATCAATGACATGCTGTCAGGTA 1900  
Qy 481 TCCGCGCGTACTCTCTGCTGCGCAGCAGCCATCAAGCGCATCTGCGCAAGTGAC 540  
Db 1901 TCCGCGCGTACTCTCTGCTGCGCAGCAGCCATCAAGCGCATCTGCGCAAGTGAC 1960  
Qy 541 TGCCCGGCTTCTGCGCGCATGCTCAATGAGCTTGGGACCATATCAATATTTGTCAGTTTC 600  
Db 1961 TGCCCGGCTTCTGCGCGCATGCTCAATGAGCTTGGGACCATATCAATATTTGTCAGTTTC 2020  
Qy 601 AACATGCTCTCAGACAGGACTTACCTGCTGGCGCGGAGACCTCTGGAATTCATG 660  
Db 2021 AACATGCTCTCAGACAGGACTTACCTGCTGGCGCGGAGACCTCTGGAATTCATG 2080  
Qy 661 AGATATTTGCTGCGCAGCAGCCCTGCGCAAGTGTGAGAAAGCAATGAGCAAGACCA 720  
Db 2081 AGATATTTGCTGCGCAGCAGCCCTGCGCAAGTGTGAGAAAGCAATGAGCAAGACCA 2140  
Qy 721 TTGTTGCGAGCTGACTCAACAAAGTTCTCACCGTCTACCGCCGCTTCTCCACCCACGA 780  
Db 2141 TTGTTGCGAGCTGACTCAACAAAGTTCTCACCGTCTACCGCCGCTTCTCCACCCACGA 2200  
Qy 781 TTAATGATGAATAGCTTGGCTGATGATCAGAGCGGCGAGCCCTCTCTCCGCAATGAC 840  
Db 2201 TTAATGATGAATAGCTTGGCTGATGATCAGAGCGGCGAGCCCTCTCTCCGCAATGAC 2260  
Qy 841 TCAGCGCCCTCCGCTGTCAGCTCTGAGACGCTAG 873  
Db 2261 TCAGCGCCCTCCGCTGTCAGCTCTGAGACGCTAG 2293

RESULT 3  
AAH68528/c  
ID AAH68528 standard; DNA; 349980 BP.  
XX AC  
XX AAH68528;  
XX  
DT 26-SEP-2001 (first entry)  
XX  
DE C glutamicum coding sequence fragment SEQ ID NO: 7063.  
XX  
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
KW organic acid synthesis; ds.  
XX  
XX Corynebacterium glutamicum.  
XX  
PN EPI108790-A2.  
XX  
PD 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.  
XX 16-DEC-1999; 99JP-00377484.  
PR 07-APR-2000; 2000JP-00159162.  
PR 03-AUG-2000; 2000JP-00280988.  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX WPI; 2001-376931/40.  
XX Novel polynucleotides derived from Coryneform bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analyzing  
PT expression profile or pattern of a gene and identifying homologous gene.  
XX  
PS Disclosure; SEQ ID NO 7063; 246pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of nucleotide and protein  
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of coryneform bacterium, measuring expression amount and analysing  
CC the expression profile or expression pattern of a gene derived from  
CC Coryneform bacterium, and identifying a homologue of a gene derived from  
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino  
CC acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a nucleic acid described  
CC in the exemplification of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the European Patent Office  
XX  
SQ Sequence 349980 BP; 80727 A; 91049 C; 97618 G; 80586 T; 0 U; 0 Other;  
Query Match 100.0%; Score 873; DB 5; Length 349980;  
Best Local Similarity 100.0%; Pred. No. 4e-245;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GTCGATAGCAGCGGTTTGGGTAAGCCCATGTTGATGGCAATCGTGTGACCTGGTTG 60  
Db 128548 GTCGATAGCAGCGGTTTGGGTAAGCCCATGTTGATGGCAATCGTGTGACCTGGTTG 128489  
Qy 61 AACCCGATGCGTATTTGGACGGCTTTGTTTATCGCGCGCTGCGCGCAATACGCG 120  
Db 128488 AACCCGATGCGTATTTGGACGGCTTTGTTTATCGCGCGCTGCGCGCAATACGCG 128429  
Qy 121 GACACCGGACGCTGATTTTCGCGCTGCGCGCTTCCGCGCAAGCTCTGATCTGTTCCG 180  
Db 128428 GACACCGGACGCTGATTTTCGCGCTGCGCGCTTCCGCGCAAGCTCTGATCTGTTCCG 128369  
Qy 181 CTGCTGGGTTTCGCGCGACGAGCATTTGTCACGCCGCTGTCACGCCCAAGCTGTGCGCG 240  
Db 128368 CTGCTGGGTTTCGCGCGACGAGCATTTGTCACGCCGCTGTCACGCCCAAGCTGTGCGCG 128309  
Qy 241 TGGATCAACGCTGCTGCTGGCAGTTGTGATGACCGCATTGGCCATCAAACTGATGTGATG 300  
Db 128308 TGGATCAACGCTGCTGCTGGCAGTTGTGATGACCGCATTGGCCATCAAACTGATGTGATG 128249  
Qy 301 GGTAGTTTTCGCGGGTTTGGATCGGTGCGCTTCCGCCAATGTTGATCGCGGCTG 360  
Db 128248 GGTAGTTTTCGCGGGTTTGGATCGGTGCGCTTCCGCCAATGTTGATCGCGGCTG 128189  
Qy 361 TGGGAATCTCATCGATCGCTCCCACTCGCGCTCAGAAATCTCCAAAGTTGTGAGTGAA 420  
Db 128188 TGGGAATCTCATCGATCGCTCCCACTCGCGCTCAGAAATCTCCAAAGTTGTGAGTGAA 128129  
Qy 421 TCAAGGCTGTTCAGCTGCTCAATGACGAGCAATCAATGACATGCTGTCAGGTA 480  
Db 128128 TCAAGGCTGTTCAGCTGCTCAATGACGAGCAATCAATGACATGCTGTCAGGTA 128069  
Qy 481 TCGCGCGCTGACTCTCTGCTGCGCAGCAGCCATCAAGCGCATCTGCGCAAGTGAC 540

Db 128068 TCCGCGCGTACTCTCTTCTGCGCAGCACCACCAATGACGCGCATCTGCGAAGTGAC 128009  
Qy 541 TCCCGCGCTTCTGCGCGAGATGTCATGAGCTTCCGACCAATATCAATATGTTTACGTTT 600  
Db 128008 TGCCGCGCTTCTGCGCGATGTCATGAGCTTCCGACCAATATCAATATGTTTACGTTT 127949  
Qy 601 ACATGCTCCTCAGACAGGACCTTACCTGCTGCGCGGAGAACCTCTGGAATTCATCG 660  
Db 127948 AACATGCTCCTCAGACAGGACCTTACCTGCTGCGCGGAGAACCTCTGGAATTCATCG 127889  
Qy 661 AGATATTTGTCCTGAGAGAGGCTTCCGCGCAAGTGTGAGAAAGCAATGAGCGCAAGACCA 720  
Db 127888 AGATATTTGTCCTGAGAGAGGCTTCCGCGCAAGTGTGAGAAAGCAATGAGCGCAAGACCA 127829  
Qy 721 TTGTTGGAGCTGAGTGCACAAAGTTCTCACCGTATGCGCGGTCTCTCCACCAAGCA 780  
Db 127828 TTGTTGGAGCTGAGTGCACAAAGTTCTCACCGTATGCGCGGTCTCTCCACCAAGCA 127769  
Qy 781 TTAATGATGGAATAGCTTGGCTGATGAATCAGAAAGGCGGAGCCCTCTCCGCAATGAAC 840  
Db 127768 TTAATGATGGAATAGCTTGGCTGATGAATCAGAAAGGCGGAGCCCTCTCCGCAATGAAC 127709  
Qy 841 TCAGCGGCTCGCTGAGCTCTGACCGTAG 873  
Db 127708 TCAGCGGCTCGCTGAGCTCTGACCGTAG 127676

RESULT 4  
AAH68419/c  
ID AAH68419 standard; DNA; 627 BP.  
XX  
AC AAH68419;  
XX  
DT 26-SEP-2001 (first entry)  
XX  
DE C glutamicum coding sequence fragment SEQ ID NO: 3454.  
XX  
DE Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
KW Coryneform bacterium; ds.  
KW organic acid synthesis; ds.  
XX  
OS Corynebacterium glutamicum.  
XX  
PN EP1108790-A2.  
XX  
PD 20-JUN-2001.  
XX  
PF 18-DEC-2000; 2000EP-00127688.  
XX  
PR 16-DEC-1999; 99JP-00377484.  
PR 07-APR-2000; 2000JP-00159162.  
PR 03-AUG-2000; 2000JP-00280988.  
XX  
PA (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX  
XX WPI: 2001-376931/40.  
DR P-PSDB; AN93200.  
XX  
XX Novel polynucleotides derived from Coryneform bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analyzing  
PT expression profile or pattern of a gene and identifying homologous gene.  
XX  
XX Claim 1; SEQ ID NO 3454; 246pp + Sequence Listing; English.  
XX  
XX The present invention provides a number of nucleotide and protein  
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of coryneform bacterium, measuring expression amount and analysing  
CC the expression profile or expression pattern of a gene derived from  
CC Coryneform bacterium, and identifying a homologue of a gene derived from  
CC coryneform bacterium. Coryneform bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a nucleic acid described  
CC in the exemplification of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the European Patent Office  
XX  
SQ Sequence 627 BP; 139 A; 159 C; 185 G; 144 T; 0 U; 0 Other;  
Query Match 65.1%; Score 568; DB 5; Length 627;  
Best Local Similarity 100.0%; Pred. No. 1.5e-156;  
Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 306 GTTTTCGCGGGTTTGAATCGGTGGCTTCGCCCAAATGTTGATGCGCGGTCTGGGA 365  
Db 627 GTTTTCGCGGGTTTGAATCGGTGGCTTCGCCCAAATGTTGATGCGCGGTCTGGGA 568  
Qy 366 AATCTCATGATCGCTCCCAATCGCGCTGAGAAATCCAAAGTTGTTGATGAATCAAG 425  
Db 567 AATCTCATGATCGCTCCCAATCGCGCTGAGAAATCCAAAGTTGTTGATGAATCAAG 508  
Qy 426 GCTGTTGTTCCAGCTGCTCAACTGACGAAGCAACCAATCAATGCACTGCTCAAGTATCCGC 485  
Db 507 GCTGTTGTTCCAGCTGCTCAACTGACGAAGCAACCAATCAATGCACTGCTCAAGTATCCGC 448  
Qy 486 GCGGTACTCTCTCTTCTGCTGCGGAGCAACCCATGCGAGGCGCATCTCGCAAGTCACTGCC 545  
Db 447 GCGGTACTCTCTCTTCTGCTGCGGAGCAACCCATGCGAGGCGCATCTCGCAAGTCACTGCC 388  
Qy 546 GCGTTCTGCGGAGTGTCTTTCAGCTTTCGAGCAATATCAATATGTTTTCAGTTCAACAT 605  
Db 387 GCGTTCTGCGGAGTGTCTTTCAGCTTTCGAGCAATATCAATATGTTTTCAGTTCAACAT 328  
Qy 606 GCGCTCAGACAGGACTTACCTTGGCTGCGGAGCAACCCCTCTGGAATTCATCGAGATA 665  
Db 327 GCGCTCAGACAGGACTTACCTTGGCTGCGGAGCAACCCCTCTGGAATTCATCGAGATA 268  
Qy 666 TTTGTCGTGAGCAGGCGCTTTCAGCTTTCGAGCAATGAGCAAGCAATGAGCAAGCAATGTT 725  
Db 267 TTTGTCGTGAGCAGGCGCTTTCAGCTTTCGAGCAATGAGCAAGCAATGAGCAAGCAATGTT 208  
Qy 726 GGCAGCTGACTGCAACAAAGTTTCTACCGCTCATCCCGCGTTCCTCCACCACCAATTAAT 785  
Db 207 GGCAGCTGACTGCAACAAAGTTTCTACCGCTCATCCCGCGTTCCTCCACCACCAATTAAT 148  
Qy 786 GATGGAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCGCCATGAACTCAGC 845  
Db 147 GATGGAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCGCCATGAACTCAGC 88  
Qy 846 GCGCTCGCTGTGAGCTCTGAGCCGTAG 873  
Db 87 GCGCTCGCTGTGAGCTCTGAGCCGTAG 60

RESULT 5  
AAH68077/c  
ID AAH68077 standard; DNA; 993 BP.  
XX  
AC AAH68077;  
XX  
DT 11-APR-2001 (first entry)  
XX  
DE Corynebacterium glutamicum MCT protein encoding DNA SEQ ID NO:669.  
XX  
KW Corynebacterium glutamicum; brevivacterium lactofermentum; MCT;  
KW membrane construction and membrane transport protein; petroleum spill;  
KW hydrocarbon degradation; gram positive aerobic bacterium; marker;  
KW identification; microorganism; fine chemical production; transformation;  
KW genome mapping; genetic engineering; ds.  
XX  
OS Corynebacterium glutamicum.  
XX  
PN WO200100805-A2.  
XX

PD 04-JAN-2001.  
 XX 23-JUN-2000; 2000WO-IB000926.  
 XX 25-JUN-1999; 99US-0141031P.  
 PR 08-JUL-1999; 99DE-01031454.  
 PR 08-JUL-1999; 99DE-01031478.  
 PR 08-JUL-1999; 99DE-01031563.  
 PR 08-JUL-1999; 99DE-01032122.  
 PR 09-JUL-1999; 99DE-01032124.  
 PR 09-JUL-1999; 99DE-01032125.  
 PR 09-JUL-1999; 99DE-01032128.  
 PR 09-JUL-1999; 99DE-01032180.  
 PR 09-JUL-1999; 99DE-01032182.  
 PR 09-JUL-1999; 99DE-01032190.  
 PR 09-JUL-1999; 99DE-01032191.  
 PR 09-JUL-1999; 99DE-01032209.  
 PR 09-JUL-1999; 99DE-01032212.  
 PR 09-JUL-1999; 99DE-01032227.  
 PR 09-JUL-1999; 99DE-01032228.  
 PR 09-JUL-1999; 99DE-01032229.  
 PR 09-JUL-1999; 99DE-01032230.  
 PR 14-JUL-1999; 99DE-01032927.  
 PR 14-JUL-1999; 99DE-01033005.  
 PR 14-JUL-1999; 99DE-01033006.  
 PR 27-AUG-1999; 99DE-01040764.  
 PR 27-AUG-1999; 99DE-01040765.  
 PR 27-AUG-1999; 99DE-01040766.  
 PR 27-AUG-1999; 99DE-01040830.  
 PR 27-AUG-1999; 99DE-01040831.  
 PR 27-AUG-1999; 99DE-01040832.  
 PR 27-AUG-1999; 99DE-01040833.  
 PR 31-AUG-1999; 99DE-01041378.  
 PR 31-AUG-1999; 99DE-01041379.  
 PR 31-AUG-1999; 99DE-01041395.  
 PR 03-SEP-1999; 99DE-01042077.  
 PR 03-SEP-1999; 99DE-01042078.  
 PR 03-SEP-1999; 99DE-01042079.  
 PR 03-SEP-1999; 99DE-01042088.  
 XX (BADI ) BASF AG.  
 PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
 PI WPI; 2001-071486/08.  
 XX P-PSDB; AAB76844.  
 PT Corynebacterium glutamicum nucleic acids encoding membrane construction  
 PT and membrane transport proteins or their portions, useful for typing or  
 PT identifying C. glutamicum or related bacteria, and as markers for  
 PT transformation.  
 XX Claim 3; Page 1107-1108; 1119pp; English.  
 XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane  
 CC construction and membrane transport (MCT) proteins given in AAB76510 to  
 CC AAB76847. The MCT nucleic acids and proteins are useful in the  
 CC identification of microorganisms which can be used to produce fine  
 CC chemicals, for modulating fine chemical production in C. glutamicum or  
 CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or  
 CC identification of C. glutamicum or related bacteria, as reference points  
 CC for mapping C. glutamicum genome, and as markers for transformation.  
 CC AAF68082 and AAF68082 represent sequencing primers which are used in an  
 CC example from the present invention  
 XX Sequence 993 BP; 222 A; 247 C; 277 G; 247 T; 0 U; 0 Other;  
 SQ  
 Query Match 46.3%; Score 404.2; DB 4; Length 993;  
 Best Local Similarity 98.8%; Pred. No. 2.7e-108;  
 Matches 418; Conservative 0; Mismatches 3; Indels 2; Gaps 1;  
 QY 453 AGCACCAATCAATGCACTGGTCAACGGTA--TCCGCGCCGCTACTCTCTTGGCTCGGCGAGC 510  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 993 AGCACCAATCAATGCACTGGTCAACGGTAATCCGCGCCGCTACTCTCTTGGCTCGGCGAGC 934  
 QY 511 ACCCATGCAAGCGCCATCTGGCGAGTGAAGTCCGCGCGCTTCTGGCGGATGTCAATGAGC 570  
 Db 933 ACCCATGCAAGCGCCATCTGGCGAGTGAAGTCCGCGCGCTTCTGGCGGATGTCAATGAGC 874  
 QY 571 TTGCGGACCATATCAATATTGTTTCACGTTCAACATGCGCCTCAGACAGGGAATACCTTGG 630  
 Db 873 TTGCGGACCATATCAATATTGTTTCACGTTCAACATGCGCCTCAGACAGGGAATACCTTGG 814  
 QY 631 CTGGCGGGGAAACCTCTCTGGAATTCATCGAGATATTTGTCGCTGAGCAGGCGCTTGGCGCA 690  
 Db 813 CTGGCGGGGAAACCTCTCTGGAATTCATCGAGATATTTGTCGCTGAGCAGGCGCTTGGCGCA 754  
 QY 691 AGTGTGTGAGAAAGCAATGAAGCAAGCAATTTGTCGCGAGTGAATGCAACAGTTCTCA 750  
 Db 753 AGTGTGTGAGAAAGCAATGAAGCAAGCAATTTGTCGCGAGTGAATGCAACAGTTCTCA 694  
 QY 751 CCGTCATCGCCCGGTTCTCTCCACCAACGATTAATGATGAATAGTTGGCTGATGAATC 810  
 Db 693 CCGTCATCGCCCGGTTCTCTCCACCAACGATTAATGATGAATAGTTGGCTGATGAATC 634  
 QY 811 AGAAGCGGCGAGCGCTCTCTCCGCAATGAGTCAAGCGCGCTCGGCTGAGCTCTGGAGCG 870  
 Db 633 AGAAGCGGCGAGCGCTCTCTCCGCAATGAGTCAAGCGCGCTCGGCTGAGCTCTGGAGCG 574  
 QY 871 TAG 873  
 Db 573 TAG 571  
 RESULT 6  
 AAF71779  
 ID AAF71779 standard; DNA; 822 BP.  
 XX AC AAF71779;  
 XX 30-APR-2001 (first entry)  
 DT DT  
 XX DE  
 XX DE  
 XX OS  
 XX KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;  
 KW fine chemical production; microorganism; organic acid; nucleoside;  
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;  
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
 KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.  
 XX PN WO200100843-A2.  
 XX PD 04-JAN-2001.  
 XX PF 23-JUN-2000; 2000WO-IB000923.  
 XX 25-JUN-1999; 99US-0141031P.  
 PR 01-JUL-1999; 99DE-01030476.  
 PR 02-JUL-1999; 99US-0142101P.  
 PR 08-JUL-1999; 99DE-01031415.  
 PR 08-JUL-1999; 99DE-01031418.  
 PR 08-JUL-1999; 99DE-01031419.  
 PR 08-JUL-1999; 99DE-01031420.  
 PR 08-JUL-1999; 99DE-01031424.  
 PR 08-JUL-1999; 99DE-01031428.  
 PR 08-JUL-1999; 99DE-01031434.  
 PR 08-JUL-1999; 99DE-01031435.  
 PR 08-JUL-1999; 99DE-01031443.  
 PR 08-JUL-1999; 99DE-01031453.  
 PR 08-JUL-1999; 99DE-01031457.  
 PR 08-JUL-1999; 99DE-01031465.  
 PR 08-JUL-1999; 99DE-01031478.  
 PR 08-JUL-1999; 99DE-01031510.  
 PR 08-JUL-1999; 99DE-01031541.



PR 08-JUL-1999; 99DE-01031573.  
 PR 08-JUL-1999; 99DE-01031592.  
 PR 08-JUL-1999; 99DE-01031632.  
 PR 08-JUL-1999; 99DE-01031634.  
 PR 08-JUL-1999; 99DE-01031636.  
 PR 09-JUL-1999; 99DE-01032125.  
 PR 09-JUL-1999; 99DE-01032126.  
 PR 09-JUL-1999; 99DE-01032130.  
 PR 09-JUL-1999; 99DE-01032186.  
 PR 09-JUL-1999; 99DE-01032206.  
 PR 09-JUL-1999; 99DE-01032227.  
 PR 09-JUL-1999; 99DE-01032228.  
 PR 09-JUL-1999; 99DE-01032229.  
 PR 09-JUL-1999; 99DE-01032230.  
 PR 14-JUL-1999; 99DE-01032922.  
 PR 14-JUL-1999; 99DE-01032926.  
 PR 14-JUL-1999; 99DE-01032928.  
 PR 14-JUL-1999; 99DE-01033004.  
 PR 14-JUL-1999; 99DE-01033005.  
 PR 14-JUL-1999; 99DE-01033006.  
 PR 12-AUG-1999; 99SUS-0148613P.  
 PR 27-AUG-1999; 99DE-01040764.  
 PR 27-AUG-1999; 99DE-01040765.  
 PR 27-AUG-1999; 99DE-01040766.  
 PR 27-AUG-1999; 99DE-01040832.  
 PR 31-AUG-1999; 99DE-01041378.  
 PR 31-AUG-1999; 99DE-01041379.  
 PR 31-AUG-1999; 99DE-01041380.  
 PR 31-AUG-1999; 99DE-01041394.  
 PR 31-AUG-1999; 99DE-01041396.  
 PR 03-SEP-1999; 99DE-01042076.  
 PR 03-SEP-1999; 99DE-01042077.  
 PR 03-SEP-1999; 99DE-01042079.  
 PR 03-SEP-1999; 99DE-01042086.  
 PR 03-SEP-1999; 99DE-01042087.  
 PR 03-SEP-1999; 99DE-01042088.  
 PR 03-SEP-1999; 99DE-01042095.  
 PR 03-SEP-1999; 99DE-01042124.  
 PR 03-SEP-1999; 99DE-01042129.  
 PR 09-MAR-2000; 2000US-0187970P.  
 XX (BADI ) BASF AG.  
 PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
 PI Kim J, Lee H, Hwang B;  
 XX WPI; 2001-137957/14.  
 DR P-PSDB; AAB79660.  
 XX Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway  
 PT proteins, useful for producing fine chemicals in microorganisms,  
 PT including organic acids, nonproteinogenic amino acids, and purine and  
 PT pyrimidine bases.  
 XX Claim 3; Page 233-234; 173pp; English.  
 XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic  
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP  
 CC nucleic acids are useful for the production of fine chemicals in  
 CC microorganisms, including organic acids, nonproteinogenic amino acids,  
 CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated  
 CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,  
 CC vitamins, cofactors, polyketides and enzymes  
 XX Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;  
 SQ  
 Query Match 37.3%; Score 326; DB 4; Length 822;  
 Best Local Similarity 100.0%; Pred. NO. 2.5e-85;  
 Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GTCCATAGCAGCGGTTTGGTAAAGCCCATGTTGATGGCATGCTGTCACCTGTTG 60  
 DB 497 GTGCATAGCAGCGGTTTGGTAAAGCCCATGTTGATGGCATGCTGTCACCTGTTG 556

QY 61 AACCCGAATGCTATTGGAAGCGTTTGTGTTATCGCGCGCGTCCGCGCAATACGCG 120  
 DB 557 AACCCGAATGCTATTGGAAGCGTTTGTGTTATCGCGCGCGTCCGCGCAATACGCG 616  
 QY 121 GACACCGGACGCTGGATTTTTCGCGCTGCGCGGCTGCGGCAAGCTGATCTGTTCCCG 180  
 DB 617 GACACCGGACGCTGGATTTTTCGCGCTGCGCGGCTGCGGCAAGCTGATCTGTTCCCG 676  
 QY 181 CTGCTGGGTTTTCGCGCGCAGCAGCATTTGTCACGCCCGCTGTCCAGCCCAAGGTGTGGCGC 240  
 DB 677 CTGCTGGGTTTTCGCGCGCAGCAGCATTTGTCACGCCCGCTGTCCAGCCCAAGGTGTGGCGC 736  
 QY 241 TGGATCAACGTCGCTGCGGCTGCTGATGACCGCATTTGGCCCATCAACTGATGTTGATG 300  
 DB 737 TGGATCAACGTCGCTGCGGCTGCTGATGACCGCATTTGGCCCATCAACTGATGTTGATG 796  
 QY 301 GGTAGTTTTCGCGCGGCTTTTGGATC 326  
 DB 797 GGTAGTTTTCGCGCGGCTTTTGGATC 822  
 RESULT 7  
 AAS96098  
 ID AAS96098 standard; DNA; 822 BP.  
 XX  
 AC AAS96098;  
 XX  
 DT 26-FEB-2002 (first entry)  
 XX  
 DE C. glutamicum gene #23 encoding metabolic pathway protein.  
 XX  
 KW Metabolic pathway protein; MP; lysine biosynthesis pathway;  
 KW methionine biosynthesis pathway; large-scale production of fine chemical;  
 KW Corynebacterium diphtheriae; diphtheria; ds.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN WO200166573-A2.  
 XX  
 PD 13-SEP-2001.  
 XX  
 PF 22-DEC-2000; 2000WO-IB002035.  
 XX  
 PR 09-MAR-2000; 2000US-0187970P.  
 PR 23-JUN-2000; 2000US-00606740.  
 XX  
 PA (BADI ) BASF AG.  
 XX  
 PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
 PI Kim J, Lee H, Hwang B;  
 XX WPI; 2001-582269/65.  
 DR P-PSDB; AAB71888.  
 XX  
 PT Nucleic acids encoding metabolic pathway proteins from Corynebacterium  
 PT glutamicum, useful for producing methionine and lysine in Corynebacterium  
 PT and Brevibacterium.  
 XX  
 PS Disclosure; Page 214-215; 316pp; English.  
 XX  
 CC The present invention relates to the isolation of novel Corynebacterium  
 CC glutamicum genes encoding metabolic pathway (MP) proteins (AAU71863-  
 CC AAU71922). The metabolic pathway proteins of the invention include  
 CC enzymes involved in the lysine and methionine biosynthetic pathways. The  
 CC polynucleotide sequences of the invention can be used for the large-scale  
 CC production and/or modulation of expression of fine chemicals such as  
 CC lysine and methionine. The sequences of the invention may be used to  
 CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a  
 CC subject to detect diphtheria. AAS96073-AA96132 represent C. glutamicum  
 CC genes encoding the novel metabolic pathway proteins of the invention  
 XX  
 SQ Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;

Query Match 37.3%; Score 326; DB 4; Length 822;  
Best Local Similarity 100.0%; Pred. No. 2.5e-85; Mismatches 0; Indels 0; Gaps 0;  
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGGTTG 60  
DB 497 GTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGGTTG 556  
QY 61 AACCCGATGCGTATTTGCGAGCGTTTGTGTTATCGCGCGCGTGGCGCCCAATACGCG 120  
DB 557 AACCCGATGCGTATTTGCGAGCGTTTGTGTTATCGCGCGCGTGGCGCCCAATACGCG 616  
QY 121 GACACCGAGCGTGATTTTCGCGCTGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 180  
DB 617 GACACCGAGCGTGATTTTCGCGCTGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 676  
QY 181 CTGCTGGGTTTCGCGCGCAGCAGCATTTGTCAAGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 240  
DB 677 CTGCTGGGTTTCGCGCGCAGCAGCATTTGTCAAGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 736  
QY 241 TGGATCAACGTCGTGCGCGAGCAGCATTTGTCAAGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 300  
DB 737 TGGATCAACGTCGTGCGCGAGCAGCATTTGTCAAGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 796  
QY 301 GGTAGTTTTCGCGCGGTTTGGCAATC 326  
DB 797 GGTAGTTTTCGCGCGGTTTGGCAATC 822

RESULT 8  
ACC80941  
ID ACC80941 standard; DNA; 711 BP.  
XX  
AC ACC80941;  
XX  
DT 27-OCT-2003 (revised)  
DT 11-AUG-2003 (first entry)  
XX  
XX LysE protein encoding sequence.  
XX L-lysine; L-arginine; LysE; ds.  
XX Corynebacterium glutamicum.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..711  
XX FT /\*tag= a  
XX FT /product= "lysE protein"  
XX  
XX EP1266966-A2.  
XX  
XX 18-DEC-2002.  
XX  
XX 05-JUN-2002; 2002EP-00012539.  
XX  
XX 12-JUN-2001; 2001JP-00177075.  
XX  
XX (AJIN ) AJINOMOTO CO INC.  
XX  
XX Gunji Y, Yasueda H;  
XX WPI; 2003-241171/24.  
XX P-PSDB; ABR58213.  
XX  
XX Novel DNA encoding variant of LysE protein from a coryneform bacterium,  
XX when introduced into methanol assimilating bacterium, facilitates  
XX excretion of L-lysine and/or L-arginine to outside of a cell.  
XX  
XX Example 1; Page 17-18; 23pp; English.  
XX  
XX The present invention relates to DNA encoding variants of protein with  
XX loop region and six hydrophobic helices which facilitates excretion of L-  
XX lysine and/or L-arginine to outside of cell of a methanol assimilating

CC bacterium when introduced into the bacterium. The method is used for  
CC encoding a protein which facilitates excretion of L-lysine, L-arginine or  
CC both of these L-amino acids to outside of a cell of a methanol  
CC assimilating bacterium when DNA of the method is introduced into the  
CC bacterium. The present sequence represents a lysE protein from  
CC Brevibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003  
XX to standardise OS field)  
SQ Sequence 711 BP; 135 A; 173 C; 222 G; 181 T; 0 U; 0 Other;  
Query Match 35.1%; Score 306; DB 8; Length 711;  
Best Local Similarity 100.0%; Pred. No. 1.8e-79;  
Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGGTTG 60  
DB 406 CTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGGTTG 455  
QY 61 AACCCGATGCGTATTTGCGAGCGTTTGTGTTATCGCGCGCGTTCGCGCAATACGCG 120  
DB 466 AACCCGATGCGTATTTGCGAGCGTTTGTGTTATCGCGCGCGTTCGCGCAATACGCG 525  
QY 121 GACACCGAGCGTGATTTTCGCGCTGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 180  
DB 526 GACACCGAGCGTGATTTTCGCGCTGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 585  
QY 181 CTGCTGGGTTTCGCGCGCAGCAGCATTTGTCAAGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 240  
DB 586 CTGCTGGGTTTCGCGCGCAGCAGCATTTGTCAAGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 645  
QY 241 TGGATCAACGTCGTGCGCGAGCAGCATTTGTCAAGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 300  
DB 646 TGGATCAACGTCGTGCGCGAGCAGCATTTGTCAAGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 705  
QY 301 GGTAG 306  
DB 706 GGTAG 711

RESULT 9  
ACC80942  
ID ACC80942 standard; DNA; 712 BP.  
XX  
AC ACC80942;  
XX  
DT 27-OCT-2003 (revised)  
DT 11-AUG-2003 (first entry)  
XX  
XX LysE24 protein encoding sequence.  
XX L-lysine; L-arginine; LysE24; ds.  
XX Corynebacterium glutamicum.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..711  
XX FT /\*tag= a  
XX FT /product= "lysE protein"  
XX  
XX EP1266966-A2.  
XX  
XX 18-DEC-2002.  
XX  
XX 05-JUN-2002; 2002EP-00012539.  
XX  
XX 12-JUN-2001; 2001JP-00177075.  
XX  
XX (AJIN ) AJINOMOTO CO INC.  
XX  
XX Gunji Y, Yasueda H;  
XX WPI; 2003-241171/24.  
XX P-PSDB; ABR58214.



XX WO200277183-A2.  
 XX  
 XX  
 XX 03-OCT-2002.  
 XX  
 XX 21-MAR-2002; 2002WO-US009107.  
 XX  
 XX 21-MAR-2001; 2001US-00815242.  
 XX 06-SEP-2001; 2001US-00948993.  
 XX 23-OCT-2001; 2001US-0342923P.  
 XX 08-FEB-2002; 2002US-00072851.  
 XX 06-MAR-2002; 2002US-0362699P.  
 XX  
 XX (ELIT-) ELITRA PHARM INC.  
 XX  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX P-PSDB; ABU25781.  
 XX  
 XX WPI; 2003-029926/02.  
 XX  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 XX for homologous nucleic acids required for cellular proliferation to  
 XX isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX Claim 14; SEQ ID NO 17521; 1766pp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 XX the 6213 antisense sequences given in the specification where expression  
 XX of the nucleic acid inhibits proliferation of a cell. Also included are:  
 XX (1) a vector comprising a promoter operably linked to the nucleic acid  
 XX encoding a polypeptide whose expression is inhibited by the antisense  
 XX nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 XX polypeptide or its fragment whose expression is inhibited by the  
 XX antisense nucleic acid; (4) an antibody capable of specifically binding  
 XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 XX proliferation or the activity of a gene in an operon required for  
 XX proliferation; (7) identifying a compound that influences the activity of  
 XX the gene product or that has an activity against a biological pathway  
 XX required for proliferation, or that inhibits cellular proliferation; (8)  
 XX identifying a gene required for cellular proliferation or the biological  
 XX pathway in which a proliferation-required gene or its gene product lies  
 XX or a gene on which the test compound that inhibits proliferation of an  
 XX organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 XX compound's activity; (11) a culture comprising strains in which the gene  
 XX product is overexpressed or underexpressed; (12) determining the extent  
 XX to which each of the strains is present in a culture or collection of  
 XX strains; or (13) identifying the target of a compound that inhibits the  
 XX proliferation of an organism. The antisense nucleic acids are useful for  
 XX identifying proteins or screening for homologous nucleic acids required  
 XX for cellular proliferation to isolate candidate molecules for rational  
 XX drug discovery programs, or for screening homologous nucleic acids  
 XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 XX *X. pneumonae* or *P. aeruginosa*. The present sequence is one of the target  
 XX prokaryotic essential genes. Note: The sequence data for this patent did  
 XX not form part of the printed specification, but was obtained in  
 XX electronic format directly from WIPO at  
 XX ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 1095 BP; 264 A; 247 C; 311 G; 273 T; 0 U; 0 Other;  
 XX  
 XX Query Match 29.0%; Score 253.6; DB 7; Length 1095;  
 XX Best Local Similarity 65.6%; Pred. No. 5.5e-64;  
 XX Matches 370; Conservative 0; Mismatches 194; Indels 0; Gaps 0;  
 XX  
 XX 310 TCGCGGTTTGAATCGGTGGCTTCCGCCAAATGTTGATCGCGGTCGTGGAAATC 369  
 XX 1094 TCGTGAACCTTCGAGGAGTAGACACCGCCCAATATTAATTCAGCGTCTTGGCAACA 1035  
 XX 370 TCATCGATCGCTCCAACTCGCGGTGAGAACTCCAGTTGTTGAGTCAATCAGGCTG 429  
 XX 1034 GTGTCAATATAATGCGTTCTTCAACGCTGAACCTCAAGATTATTAAGAGCTCCAGATT 975

QY 430 TTGTCACGCTGCTCAACTGACGAGACCAATCAATCAATGTCAGTGTCAAGTATCCGCGCG 489  
 DB 974 TGATCCAACTGCGCTACGAGATGCTCCGATCAAGCACTGTTACCGTGTGGCCCCA 915  
 QY 490 TACTCTCTTGTGCGGACGACCCATGCAAGCGCATCTCGCAAGTACTGCGCGCGGT 549  
 DB 914 TAAATCGCTTGTTCGCGACGACCCAGCAATAGCCAGCTGTGCCAACGTTTGGCCAGT 855  
 QY 550 TCTTGGGCGATGTCATTCAGCTTGGGACCATATCAATATTTGTTCAAGTTCACATGCC 609  
 DB 854 COTAAACGCTAAATCATTCAGTGCAGCACCATGTCAGATTTTGTAGCGTTGAGCATCTCT 795  
 QY 610 TCAGACAGGAGCTTACCTGCTGCGCGGGAACCTCTCGGAATTCATCGAGATTTTG 669  
 DB 794 TTGCCCAAGGACTTCTCTGCTGCTGCGCGTGAATCACTGCGACGCGCTCGAGATACCTG 735  
 QY 670 TCCGTGAGCAGCGCTTCCGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCATTTGTTGCA 729  
 DB 734 TCAGTTAAAGGCCCTGTGCAAGTGGGAAATGCAATTAACCAACCCAGTCGGTTATTGCGC 675  
 QY 730 GCTGACTGCAACAAGTTCTCAACGCTCATCGCGGTTCCTCCACCCAAAGATTATGATG 789  
 DB 674 GCGACTCTAACAGGTTCTCGGCACTTCTACCTGTTCTTCTACCCAGCGGTTCAAGATA 615  
 QY 790 GAATAGCTTGGCTGATGAATCAGAAAGCGGCGGCGCTCTCCGCCAATGAATCAGCGCGC 849  
 DB 614 GAATAGCTCGGTTGATGAATCAGCAGCGCGGCGGCGCTCAACAGCTAGAAATTTCCGAGATC 555  
 QY 850 TCCGCTGTGAGCTCTCGACCGTAG 873  
 DB 554 TCTCGGTTAGTTTCAGGTTCCATAG 531  
 RESULT 12  
 ID AAH45375 standard; DNA; 1568 BP.  
 AC AAH45375;  
 DT 11-SEP-2001 (first entry)  
 DE C. thermoaminoogenes lysin biosynthetic enzyme lyse DNA.  
 KW Heat-resistant; lysin biosynthesis; enzyme; coryneform;  
 KW aspartate-semialdehyde dehydrogenase; lyse; ds.  
 OS Corynebacterium thermoaminoogenes.  
 PN JP2001120270-A.  
 PD 08-MAY-2001.  
 PP 01-NOV-1999; 99JP-003111148.  
 PR 01-NOV-1999; 99JP-003111148.  
 PA (AJIN) AJINOMOTO KK.  
 DR WPI; 2001-364750/38.  
 DR P-PSDB; AA64047.  
 PT A heat-resistant lysin biosynthetic system enzyme gene of a high  
 PT temperature-resistant coryneform microbe.  
 PS Example 5; Page 22-24; 27pp; Japanese.  
 CC The invention relates to a gene from a high temperature-resistant  
 CC coryneform microbe that encodes a heat-resistant lysin biosynthetic  
 CC enzyme. The enzyme has aspartate-semialdehyde dehydrogenase activity and  
 CC can be used for growing amino acid-producing microbes. The present  
 CC sequence encodes an enzyme of the invention  
 XX Sequence 1568 BP; 258 A; 525 C; 490 G; 295 T; 0 U; 0 Other;  
 SQ



AC ACA26520;  
 XX 19-JUN-2003 (first entry)  
 XX Prokaryotic essential gene #8177.  
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 XX Burkholderia mallei.  
 XX WO200277183-A2.  
 XX 03-OCT-2002.  
 XX 21-MAR-2002; 2002WO-US009107.  
 XX 21-MAR-2001; 2001US-00815242.  
 XX 06-SEP-2001; 2001US-00948993.  
 XX 25-OCT-2001; 2001US-0342923P.  
 XX 08-FEB-2002; 2002US-00072851.  
 XX 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 XX P-PSDB; ABU22650.  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX Claim 14; SEQ ID NO 14390; 1766pp; English.  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: the sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 1107 BP; 204 A; 360 G; 173 T; 0 U; 0 Other;  
 XX Query Match 11.1%; Score 96.8; DB 7; Length 1107;  
 XX Best Local Similarity 56.9%; Pred. No. 7.3e-18;

Matches 232; Conservative 0; Mismatches 152; Indels 24; Gaps 2;  
 QY 327 GGTGGCTTCGGCCCAATGTTGATGCGCGCTGCTGGGAATCTCATGATCGCTCCAA 386  
 DB 1092 GGACGGCTTTTCCACAGATGATCCCGCTTCGGTGGCATAGCGGTGATCTCGGGAT 1033  
 QY 387 CTGGGCTCAGAAAACCTCCAAGTGTGAGTGAATCAAGGCTGTTTCCAGCTGCTCAAC 446  
 DB 1032 CTGGTGGCTGAACGGAGATGGCGGAGCGCGGACGTTCTCGCGCACCTGCTCCGC 973  
 QY 447 TGACGAAGACCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 506  
 DB 972 GCGGCTCGCACCGATCAACGCGGACGTCAGC-----CGGAATCGCG 931  
 QY 507 CAGCACCCATGCAAGCGCATCTGCGCAAGTGAATGCTGCGCGCTTCTGGGGATGTCATT 566  
 DB 930 CAGCACCCAGGCAAGCGCATCTGCGCGGAGGCTTGGCGCGCGCTTGGCGATCTGTT 871  
 QY 567 GAGCTTGGCGGACCATATCAATATTTTCACTGTTCAACATGCTCCAGACGAGACTTACC 626  
 DB 870 CAGCTTGGCGACGTCGTCGAGGTTCTCGCGCTCAGGTGAGCTTCTTTCAGCGATCGCC 811  
 QY 627 -----CTGGCTGGCGCGGGAACCTCTGGAATTCATGAGATATTTGTCGAGGAGCAG 680  
 DB 810 GCGCGGCTTGTGATCCCGCATCCGCGGACGCGCTTCAGATATCTGAGGTCAGAG 751  
 QY 681 GCGCTGGCGAAGTGTGAGAAAGCAATGACGCAAGACCAATGTTGTCGC 728  
 DB 750 CCGCTGGCGAGCGGCTGAACCGGATGACGCGCGCGCTTCTGCTC 703  
 RESULT 15  
 ACA24017/c  
 ID ACA24017 standard; DNA; 1041 BP.  
 XX ACA24017;  
 XX 19-JUN-2003 (first entry)  
 XX Prokaryotic essential gene #5674.  
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 XX Borrelia cepacia.  
 XX WO200277183-A2.  
 XX 03-OCT-2002.  
 XX 21-MAR-2002; 2002WO-US009107.  
 XX 21-MAR-2001; 2001US-00815242.  
 XX 06-SEP-2001; 2001US-00948993.  
 XX 25-OCT-2001; 2001US-0342923P.  
 XX 08-FEB-2002; 2002US-00072851.  
 XX 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 XX P-PSDB; ABU20147.  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX Claim 14; SEQ ID NO 11887; 1766pp; English.  
 XX The invention relates to an isolated nucleic acid comprising any one of

Thu Mar 18 12:31:02 2004

CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: the sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1041 BP; 193 A; 338 C; 350 G; 160 T; 0 U; 0 Other;

Query Match	11.0%	Score 96.2;	DB 7;	Length 1041;
Best Local Similarity	55.7%	Pred. NO. 1.1e-17;		
Matches 241;	Conservative 0;	Mismatches 168;	Indels 24;	Gaps 2;

  

QY	327	GGTGGCTTCGCGCAATGTTGATCGCGCGTGTGGAAATCTCATCGATCGCTCCCAA	386
DB	1026	GSAGCGCTTTCCACAGATTGATGCGCCCTTTCGTCGGTAACTGATCTCCCGAG	967
QY	387	CTCGGCGTCAGAAATCCAAAGTTGTGAGTGAATCAAGGCTGTTGTCCAGCTGCTCAAC	446
DB	966	TTCTCTCGCGAGAAATTCAGAGTTCGCAACGACGAGCTTCTCGCGCACCTGCTCGGC	907
QY	447	TGACGAAGCACCAATCAATCACTGTGTACGCTATCGCGCGCTACTCTCTGTCGCG	506
DB	906	CCGCGCTCGCGCGATCAGCGCGAGGTCAACGCGCGG-----TTACG	865
QY	507	CAGCAGCATGCAAGCGCCATCTGCGAGTGAATGCTCGCGCGCTTCTGGCGATGTCATT	566
DB	864	CAGCAGCATGCAAGCGCCATCTGCGAGTGAATGCTCGCGCGCTTCTGGCGATGTCATT	805
QY	567	GAGCTTGGCGACCATATCAATTTGTTACGTTCAATGCGCTTCAACATGCGCTCAGACAGGACTTACC	626
DB	804	GAGCTTGGCGACCATATCAATTTGTTACGTTCAATGCGCTTCAACATGCGCTCAGACAGGACTTACC	745
QY	627	-----CTGGCTGGCGCGGACCCCTCTGGATTCATCGAGATATTTGTCGTCGAGCAG	680
DB	744	GCCCGGCTTGTTCACGCGCGCTCGCGCGCACCGCTTCAGGTACTTCGACGTCGAGCAA	685
QY	681	GCCCTGGCGAGTGTGAGAAAGCAATGACGCAAGACCATTTGTTGGCAGCTGACTGCAA	740
DB	684	GCCCTGGCGAGTGTGAGAAAGCAATGACGCAAGACCATTTGTTGGCAGCTGACTGCAA	740
QY	741	CAAGTCTCTACCG	753
DB	624	CAGATCTCTTCG	612

Search completed: March 15, 2004, 14:59:16  
 Job time : 357.983 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:47:35 ; Search time 65.7645 Seconds  
(without alignments)  
7366.769 Million cell updates/sec

Title: US-09-105-117K-1\_COPY\_1421\_2293  
Perfect score: 873  
Sequence: 1 gtgcgaacgacgggttg.....ctgfgagctctggaccgtag 873

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71.6	8.2	5541	1 US-08-920-812-20	Sequence 20, Appl
2	71.6	8.2	5541	1 US-08-920-827-20	Sequence 20, Appl
3	71.6	8.2	5541	1 US-08-921-177-20	Sequence 20, Appl
4	71.6	8.2	5541	1 US-08-362-577C-20	Sequence 20, Appl
5	71.6	8.2	5541	2 US-08-920-828-20	Sequence 20, Appl
6	61.8	7.1	699	4 US-09-252-991A-7908	Sequence 7908, Ap
7	61.8	7.1	834	4 US-09-252-991A-7643	Sequence 7643, Ap
8	61.8	7.1	894	4 US-09-252-991A-7841	Sequence 7841, Ap
9	61.2	7.0	1095	4 US-09-489-039A-5370	Sequence 5370, Ap
10	57	6.5	597	4 US-09-894-844-11	Sequence 11, Appl
11	57	6.5	15239	3 US-08-390-878-17	Sequence 17, Appl
12	57	6.5	4403765	3 US-09-103-840A-2	Sequence 2, Appl
13	57	6.5	4411529	3 US-09-103-840A-1	Sequence 1, Appl
14	53.8	6.2	1176	4 US-09-252-991A-7571	Sequence 7571, Ap
15	50.2	5.8	1002	4 US-09-724-623-24	Sequence 24, Appl
16	45	5.2	642	4 US-09-489-039A-4674	Sequence 4674, Ap
17	39.6	4.5	1935	2 US-08-492-027A-9	Sequence 9, Appl
18	37.6	4.3	4403765	3 US-09-103-840A-2	Sequence 2, Appl
19	37.6	4.3	4411529	3 US-09-103-840A-1	Sequence 1, Appl
20	36.2	4.1	1233	4 US-09-252-991A-2121	Sequence 2121, Ap
21	36.2	4.1	1395	4 US-09-252-991A-1954	Sequence 1954, Ap
22	36.2	4.1	1578	4 US-09-489-039A-3846	Sequence 3846, Ap
23	36.2	4.1	4056	4 US-09-489-039A-3861	Sequence 3861, Ap
24	35.4	4.1	1077	4 US-09-252-991A-14964	Sequence 14964, A
25	35.4	4.1	1155	4 US-09-252-991A-14787	Sequence 14787, A
26	35	4.0	1002	4 US-09-252-991A-1099	Sequence 1099, Ap
27	35	4.0	1092	4 US-09-252-991A-946	Sequence 946, Appl

C 28	35	4.0	1521	4 US-09-252-991A-989	Sequence 989, App
C 29	34.8	4.0	2730	4 US-09-489-039A-6744	Sequence 6744, Ap
C 30	34.6	4.0	417	4 US-09-252-991A-1351	Sequence 1351, Ap
C 31	34.6	4.0	1029	4 US-09-252-991A-1752	Sequence 1752, Ap
C 32	34.6	4.0	1098	4 US-09-252-991A-1304	Sequence 1304, Ap
C 33	34.6	4.0	1215	4 US-09-252-991A-1975	Sequence 1975, Ap
C 34	34.6	4.0	1230	4 US-09-252-991A-2136	Sequence 2136, Ap
C 35	34.6	4.0	1632	4 US-09-252-991A-1216	Sequence 1216, Ap
C 36	34.6	4.0	1746	4 US-09-252-991A-1163	Sequence 1163, Ap
C 37	34.6	4.0	1995	4 US-09-252-991A-14545	Sequence 14545, A
C 38	34.6	4.0	3396	4 US-09-252-991A-14676	Sequence 14676, A
C 39	34.6	4.0	3444	4 US-09-252-991A-15078	Sequence 15078, A
C 40	34.4	3.9	513	4 US-09-252-991A-2092	Sequence 2092, Ap
C 41	34.4	3.9	1500	4 US-09-252-991A-1868	Sequence 1868, Ap
C 42	34.4	3.9	1737	4 US-09-252-991A-1717	Sequence 1717, Ap
C 43	34.4	3.9	1767	4 US-09-252-991A-2007	Sequence 2007, Ap
C 44	34.4	3.9	2048	3 US-08-776-246-1	Sequence 1, Appli
C 45	34.4	3.9	2048	3 US-08-776-251-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-08-920-812-20  
; Sequence 20, Application US/08920812  
; Patent No. 5763188  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matshisa, Akio  
; APPLICANT: Uehara, Hirotsugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,812  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,577  
; FILING DATE: 27-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 19036/32420  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5541 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli  
; STRAIN: Clinical Isolate BC-625  
; US-08-920-812-20

Query Match 8.2%; Score 71.6; DB 1; Length 5541;  
Best Local Similarity 49.4%; Pred. No. 2.8e-11;  
Matches 337; Conservative 0; Mismatches 299; Indels 46; Gaps 4;

203 CATTTGTCAGCCCGCTGCCAGCCCAAGGTGTGGCTGGATCAACGTCGTGGTGGCAG 262  
Db CATATTCATCTGACGTTGTGCTCTCTGATTTGTATGGCTATCTTACCGACTGAC 2746

263 TTGTGATGACCGAATGGCCATCAAACTGATGTTGATGGTGTAGTTTCGGGGTTTGG 322  
Db TGTCTATAATCACTGATCTTATCGGGCGGGTCTCGCGCCGCTTATTAACAGTCAAT 2806

323 ATTCGGTGGCTTGGCCAAATGTTGATGCCGGTGTGGGGAATCTCATCGCTGCT 382  
Db TATCGGAAGAGCGCTGCCAGATTGATGCTGCCATCGCGGATATGCTGATCAATCTGG 2866

383 CCAACTCGGGGTGAGAAACTCCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGACTGCT 442  
Db CCAGCTCTCGGTGCTAAATGTCAGATTTATTCAGCGCTGCAAGTCTCTCAAGTTGTC 2926

443 CAACTGACGAAGCAACAAATCAATGATCTGATGATCGGTATCGGCGGTACTCTCTGCT 502  
Db CGCGCGGC-TGGCAACCAATCAATACCGACGTCAACGATCA-----T 2967

503 CGCGCAGCAGCCATGCAAGCGCATCTGCGCAAGTCACTGCCCGCTTCTCGGGAGT 562  
Db CTTTCAGCAACAGCTTAACGCCATTTGCGCCATTTGATTTGCTCCAGCTGTGCGCATTT 3027

563 CATTGAGCTTGGCAACATATCAATATTTTACGTTTCAATGCTTCAATGCTTCAACAGGACT 622  
Db CATTCAATAAGTGTAGGCTTGTGAGTTGGCTTGGTAAAGCAATTCAGCGCTTCAAGTTC 3087

623 TACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 670  
Db GAACTTTATTCCTTCAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3147

671 CGGTGAGCAGCCCTGCGCAAGTGTGAGAAAGCAATGAGCAAGCAATGTTGGCGAG 730  
Db CGGTGAGCAATCTTTCAGCAGAGGATTAAGCAATATACAGCCCAAGCGCTTATTTGCA 3207

731 CTGACTGCAACAGTCTTCAACGCTATCGCCCGGTTCTTCCACCAAGCAATTAATGATGG 790  
Db GGGTATCCAGCA-----GGCGGCTTTATTCACCCAGCGCTTCAAGTAAAT 3252

791 AATAGTGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 850  
Db TGTACGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3253

851 CGCGTGTGAGCTCTGGACGTA 872  
Db TTTGGTCTGCTTGGCGAGTA 3334

RESULT 2  
US-08-920-827-20  
; Sequence 20, Application US/08920827  
; Patent No. 5770375  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Teuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotsugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/920,827  
FILING DATE: 29-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,577  
FILING DATE: 27-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 19036/32420  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5541 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORGANISM: Escherichia coli  
STRAIN: Clinical Isolate EC-625  
US-08-920-827-20

Query Match 8.2%; Score 71.6; DB 1; Length 5541;  
Best Local Similarity 49.4%; Pred. No. 2.8e-11;  
Matches 337; Conservative 0; Mismatches 299; Indels 46; Gaps 4;

203 CATTGTGACCGCCCTGTCCAGCCCAAGGTGTGGCTGGATCAACGTCGTGGTGGCAG 262  
Db CATTATCCATCTGACGTTGTGCTCTCTGCAATTTGTGATGGGCTATCTTACCGACTGAC 2746

263 TTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGTGTAGTTTCGGGGTTTGG 322  
Db TGTCTATAATCACTGATCTTATCGGGCGGGTCTCGCGCCGCTTATTAACAGTCAAT 2806

323 AATCGTGGCTTTCGCCAAATGTTGATGCCGGTGTGGGGAATCTCATCGCTGCT 382  
Db TATCGGAAGAGCGCTGCCAGATTGATGCTGCCATCGCGGATATGCTGATCAATCTGG 2866

383 CCAACTCGCGCTGAGAAACTCCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGACTGCT 442  
Db CCAGCTCTCGGTGCTAAATGTCAGATTTATTCAGCGCTTCCAGCTTCTCTCTTGGT 2926

443 CAACTGACGAAGCAACCAATCAATGCACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 502  
Db TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2967

503 CGCGCAGCAGCCATGCAAGCGCATCTGCGCAAGTCACTGCCCGCTTCTCGGGAGT 562  
Db CTTTCAGCAACAGCTTAACGCCATTTGCGCCATTTGATTTGCTCCAGCTGTGCGCATTT 3027

563 CATTGAGCTTGGCAACATATCAATATTTTACGTTTCAACATGCTTCAACAGGACT 622  
Db CATTCAATAAGTGTAGGCTTGTGAGTTGGCTTGGTAAAGCAATTCAGCGCTTCAAGTTC 3087

623 TACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 670  
Db GAACTTTATTCCTTCAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3147

671 CGGTGAGCAGCCCTTGGCAAGTGTGAGAAAGCAATGAGCAAGCAATGTTGGCGAG 730  
Db CGGTGAGCAATCTTTCAGCAGAGGATTAAGCAATATACAGCCCAAGCGCTTATTTGCA 3207

731 CTGACTGCAACAGTCTTCAACGCTATCGCCCGGTTCTTCCACCAAGCAATTAATGATGG 790  
Db GGGTATCCAGCA-----GGCGGCTTTTATTCACCCAGCGCTTCAAGTAAAT 3252

791 AATAGTGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 850  
Db TGTACGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3253

851 CGCGTGTGAGCTCTGGACGTA 872  
Db TTTGGTCTGCTTGGCGAGTA 3334

QY 791 AATAGCTTGGCTGATGATCAGAGCGGCGAGCCCTCTCCGCGCATGAACTCAGCGCGCT 850  
Db 3253 TGTACAGAGGTTGATGAATTAACAGCGGAATTTTCCACTCGCGCAGCACTCAACCACTTT 3312  
QY 851 CCGCTGTGAGCTCTCGACCGTA 872  
Db 3313 TTTGCGTCCGCTCTGGCGAGTA 3334

RESULT 3  
US-08-921-177-20  
; Sequence 20, Application US/08921177  
; Patent No. 5798211  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotsugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/921,177  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,577  
; FILING DATE: 27-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 19036/32420  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5541 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; ORGANISM: Escherichia coli  
; STAIN: Clinical Isolate EC-625  
US-08-921-177-20

Query Match 8.2%; Score 71.6; DB 1; Length 5541;  
Best Local Similarity 49.4%; Pred. No. 2.8e-11;  
Matches 337; Conservative 0; Mismatches 299; Indels 46; Gaps 4;

QY 203 CATTTGATCGCGCGCTGTCAGCGCCGAGGTGGCGTGGATCAACGTCGTCGTCGAG 262  
Db 2687 CATTATCCATCTGACGCTTTGTGCTCTCTGCAATTTGATGGGCTATCTTGACCGACTGAC 2746  
QY 263 TTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGTGTAGTTTTCGCGGTTTGG 322  
Db 2747 TGTCTATATCATGATCTATGTCGGGCGCGGTTCTCGCGCCGCTTATTAACAGTCATT 2806  
QY 323 AATCGGTGGCTTCGCCCCAAATGTTGATCGCGCGCTGCTGGGAAATCTCATGATCGCCT 382

Db 2807 TATCGAAGACGCGCTGCCACAGATTGAGTTCGCCATCGGCGATATGCTGATCAATCTGG 2866  
QY 383 CCAACTCGGCGTCAAGAACTCCAGATTGTTGAGTGAATCAAGCTGTTGTCAGCTGCT 442  
Db 2867 CAGCTCTCGGTGCTAAATGTCAGATTATTCAGCGCTGACGCTTCTCTCAAGTTGTC 2926  
QY 443 CAACTGACGAAGCAACCAATCAATGCACTGTCACGGTATCGGCGCGGTACTCTCTCTGCT 502  
Db 2927 CGCGGGC-TGGCAACCAATCAATACCGACGTCACCGGATCA-----T 2967  
QY 503 CGCGGACGCCCATGCAAGCGCCATCTGGCGAAGTGAAGTACTGCCCGCGGTTCTTGGGGAAT 562  
Db 2968 CTTTCAGCAACACAGCTTAAACGCCAATTTGGGCCAATTTGATTTGTCACGCTGCTGTGCCATTT 3027  
QY 563 CATTGAGCTTGGGACCATATCAATATTTGTTCACTGTTCAACATGCGCTCAGACGCGACT 622  
Db 3028 CATTCAATAAGTGTAGGCTGTGAGGTTGGCTTCGGTAAGCATTTTGGCGGTGACACAC 3087  
QY 623 TACCTCGCTGGC-----GCGGAAACCTCTGGAATTCATCGAGATATTGT 670  
Db 3088 GAACTTTATTCCTTCACGATGCACTCGTGAATCTTTGGGAATGCGGTTGAGATATTTTC 3147  
QY 671 CCGTGAGAGCGCTTGGCAAGTGTGAGAAAGCAATGACGCCAGACCAATTTGTTGGCAG 730  
Db 3148 CGGTCAAGCAATCTCGAGCCAGAGGAGTAAGGCAATACAGCCACGCGCTTATTTTSCA 3207  
QY 731 CTGACTGCAAGAGTTCTCACCGTCATCGCGCGGTTCTCTCCACCACGATTAAATGATGG 790  
Db 3208 GGGTATCCAGCA-----GGCGGTTTATCCACCAGCGGTTCACTAAAT 3252  
QY 791 AATAGCTTGGCTGATGAATCAGAGCGGCGAGCGCTCTCTCGCCATGAACCTCAGCGGCT 850  
Db 3253 TGTACAGAGGTTGATGAATTAACAGCGGAATTTTCCACTCGCGCAGCACTCAACCACTT 3312  
QY 851 CCGCTGTGAGCTCTGACCGTA 872  
Db 3313 TTTGCGTCCGCTCTGCGGAGTA 3334

RESULT 4  
US-08-362-577C-20  
; Sequence 20, Application US/08362577C  
; Patent No. 5807673  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotsugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,577C  
; FILING DATE: 27-MAR-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 19036/32420  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5541 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORGANISM: Escherichia coli  
STRAIN: Clinical Isolate EC-625  
US-08-362-577C-20

Query Match 8.2%; Score 71.6; DB 1; Length 5541;  
Best Local Similarity 49.4%; Pred. No. 2.8e-11;  
Matches 337; Conservative 0; Mismatches 299; Indels 46; Gaps 4;  
QY 203 CATTGTGACGCGCGCTGTCAGCCCAAGGTGTGGCGTGGATCAAGCTGCTGCGGAG 262  
Db CATTATCCATCTGACGCTTGTGCTCTGCAATTTGTGATGGCTATCTTGACCGACTGAC 2746  
QY 263 TTGTGATGACCGCATTTGGCCATCAAACTGATTTGATTTGATTTGCGGGTTTGG 322  
Db TCGTCAATCACTGATCTTATGCGGGCGGTTCTCGCGCCGTTATTAACAGGTCATT 2806  
QY 323 AATCGGTGCGCTTGCCTCAATGTTGATGCGGGCTGCGGGAATCTCATCGATCGCT 382  
Db TATCGGAAGACGCTGCCACAGTTCAGCTTGCATTCGGCATATGCTGATGATCACTTCGG 2866  
QY 383 CCAACTCGCGCTCAGAAAACCTCAAGTTTGTGATGTAATCAAGGCTGTTTCCAGCTGCT 442  
Db CCAGCTCTCGTCTAATGTCAGATTTATTCAGCGCTGACGCTTCTCTCAAGTTGTC 2926  
QY 443 CAACTGACGAAGACCAATCAATGATGCTGCTGATGCTGATGCTGCTGCTGCTGCT 502  
Db CGCGCGGC-TGGCACCACCAATATACCGACGTCACGCGATCA-----T 2967  
QY 503 CGCGCAGCACCATGCAAGCGCATCTGCGCAAGTCACTGCGCGCTTCTCGGGGAGTGT 562  
Db CTTTACGACACGCTTAAGCGCATTTGCGCCATTGATTTGTCAGGCTGCTGTCGCAATT 3027  
QY 563 CATTGAGCTTCGGACCATATCAATTTATTTACGTTCAAGTTCACATGCGCTCAGACGAGCT 622  
Db CATTCAATAAGTGTAGGCTGTTGAGTTGGCTTCCGTAAGCATTTTTCGGGCTCAGACAC 3087  
QY 623 TACCTGGCTGGC-----GCGGNAACCTCTGGAATTCATTCAGATTTTGT 670  
Db GAACCTTTATTCCTTCAAGATGATCCGTTGATTTGCGGAATGCGCTGAGATATTTTC 3147  
QY 671 CCGTGACGAGCGCTCGCGCAAGTGTGAGAAAGCATGACGCAAGACCATTTGTCGAG 730  
Db CGGTGACATTCCTTGACGAGAGGTAAGGCAATACAGCCACGCGTTATTTGCA 3207  
QY 731 CTGACTGCAACAAGTCTTACCGTCAATCGCCGGTTCTCCACCCACAGATTAATGAGG 790  
Db GGCTATCCAGCA-----GCGCGCTTTATCCACCCAGCGGTTCACTAAAT 3252  
QY 791 AATAGTTGGCTGATGAATCAGAGCGGCGAGCCCTCTCCGCAATGATCACTGACGCGCT 850  
Db TGACGAAGTTGATGAATTAACAGCGGAATTTTTCACCTCGCGGAGCACTCAACATTT 3312  
QY 851 CCGCTGTGAGCTCTGAGCCGTA 872  
Db TTTGCGTCCGCTCTGGCGAGTA 3334

RESULT 5  
US-08-920-828-20  
Sequence 20, Application US/08920828  
Patent No. 5853998  
GENERAL INFORMATION:  
APPLICANT: Ohno, Tsuneya

APPLICANT: Matsuhisa, Akio  
APPLICANT: Uehara, Hirotsugu  
APPLICANT: Eda, Soji  
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/920,828  
FILING DATE: 29-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,577  
FILING DATE: 27-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 19036/32420  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5541 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli  
STRAIN: Clinical Isolate EC-625  
US-08-920-828-20

Query Match 8.2%; Score 71.6; DB 2; Length 5541;  
Best Local Similarity 49.4%; Pred. No. 2.8e-11;  
Matches 337; Conservative 0; Mismatches 299; Indels 46; Gaps 4;  
QY 203 CATTGTGACGCGCGCTGTCAGCCCAAGGTGTGGCGTGGATCAAGCTGCTGCGGAG 262  
Db CATTATCCATCTGACGCTTGTGCTCTTGTGATTTGATTTGATTTGCGGGTTTGG 322  
QY 263 TTGTGATGACCGCATTTGGCCATCAAACTGATTTGATTTGATTTGCGGGTTTGG 322  
Db TCGTCAATCACTGATCTTATGCGGGCGGTTCTCGCGCCGTTATTAACAGGTCATT 2806  
QY 323 AATCGGTGCGCTTGCCTCAATGTTGATGCGGGCTGCGGGAATCTCATCGATCGCT 382  
Db TATCGGAAGACGCTTGCCTCAGCATTCAGCTCGCATCGCGATATGCTGATCAATTCGG 2866  
QY 383 CCAACTCGCGCTCAGAAAACCTCCAAGTTTGTGATGTAATCAAGGCTGTTTCCAGCTGCT 442  
Db CCAGCTCTCGTCTAATGTCAGATTTATTCAGCGCTGACGCTTCTCTCAAGTTGTC 2926  
QY 443 CAACTGACGAAGACCAATCAATGATGCTGCTGATGCTGATGCTGCTGCTGCTGCT 502  
Db CGCGCGGC-TGGCACCACCAATCAATACCGACGTCACGCGATCA-----T 2967  
QY 503 CGCGCAGCACCATGCAAGCGCATCTGCGCAAGTCACTGCGCGCTTCTCGGGGAGTGT 562  
Db CTTTACGACACGCTTAAGCGCATTTGCGCCATTGATTTGTCAGGCTGCTGTCGCAATT 3027  
QY 563 CATTGAGCTTCGGACCATATCAATTTATTTACGTTCAAGTTCACATGCGCTCAGACGAGCT 622  
Db CATTCAATAAGTGTAGGCTGTTGAGTTGGCTTCCGTAAGCATTTTTCGGGCTCAGACAC 3087  
QY 623 TACCTGGCTGGC-----GCGGNAACCTCTGGAATTCATTCAGATTTTGT 670  
Db GAACCTTTATTCCTTCAAGATGATCCGTTGATTTGCGGAATGCGCTGAGATATTTTC 3147  
QY 671 CCGTGACGAGCGCTCGCGCAAGTGTGAGAAAGCATGACGCAAGACCATTTGTCGAG 730  
Db CGGTGACATTCCTTGACGAGAGTAAAGCAATACAGCCACGCGTTATTTGCA 3207  
QY 731 CTGACTGCAACAAGTCTTACCGTCAATCGCCGGTTCTCCACCCACAGATTAATGAGG 790  
Db GGCTATCCAGCA-----GCGCGCTTTATCCACCCAGCGGTTCACTAAAT 3252  
QY 791 AATAGTTGGCTGATGAATCAGAGCGGCGAGCCCTCTCCGCAATGATCACTGACGCGCT 850  
Db TGACGAAGTTGATGAATTAACAGCGGAATTTTTCACCTCGCGGAGCACTCAACATTT 3312  
QY 851 CCGCTGTGAGCTCTGAGCCGTA 872  
Db TTTGCGTCCGCTCTGGCGAGTA 3334

Db 3028 CATTCAATAGTGTAGCTGTGAGCTTGGCTTCGGTAAGCAATTTTGGCGTCAGACAC 3087  
Qy 623 TACCTGGCTGGC-----GCGGAACCTCTGGAATTCATCAGATATTTGT 670  
Db 3088 GAACCTTTATCCCTTCACGATGATCCGTAATCTTGGGAATGCGGTGAGATATTTTC 3147  
Qy 671 CCGTGAGAGCGCTCGCAAGTGTGAGAAGCAATGACGCAAGACATTTGTGGCAG 730  
Db 3148 CGTTCAGCAATCCCTGAGCAGAGGATAGGCAATACAGCCAGCGCTTATTTGCA 3207  
Qy 731 CTGATGCAACAAGTTCTCACCGTCAATCGCCGGTCTCTCCACCCCAAGCAATTAATGATGG 790  
Db 3208 GGGTATCCAGCA-----GCGCGCTTTTATCCACCCAGCGGTTCAATAAT 3252  
Qy 791 AATAGTCTGGCTGATGATCAGAGCGGCGAGCCCTCTCCGCAATGAACTCAGCGCCT 850  
Db 3253 TGTACGAAGTGTGATGAATTAACAGCGGAATTTTCCACTCGCGCAGCAACTCAACCATTT 3312  
Qy 851 CCGCTGTGAGCTCTGACCGTA 872  
Db 3313 TTGCGTCTGCTCTGGCGAGTA 3334

RESULT 6  
US-09-252-991A-7908  
; Sequence 7908, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7908  
; LENGTH: 699  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7908

Query Match 7.1%; Score 61.8; DB 4; Length 699;  
Best Local Similarity 54.6%; Pred. No. 8.9e-09;  
Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;  
Qy 35 TGATGGCAATCGTGTGACCTGTGACCTGTGACCCGAATGCGGTATTTGGACGCGTTTGTGTTTA 94  
Db 434 TGGCGGCACTGGCGGTACCTGCTCAACCCCACTCTATCTCGATACCGTACTTCTGA 493  
Qy 95 TCGCGGCGTTCGCGCGCAATACGCGGCAACCGGACGCGTGGATTTTCGCGCGCTGGCGGT 154  
Db 494 TCGGTTCGCTCGCGCGCCAGCAGG-----CGCGCGCGCGCGCGCTATGCGCTCGCGCGCG 547  
Qy 155 TCGCGGCAAGCTGATCTGCTTCCGCTGGTGGTTCGCGCGCAGCAGCATTTGTCAAGCC 214  
Db 548 CACGCGCTCGCTGATGTTCTTCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 607  
Qy 215 CGCTGTCCAGCCCAAGTGTGGCGCTGATCAACGCTGCGCGCGCAGCATTTGTCAAGCC 274  
Db 608 GGCTGGCG 667  
Qy 275 CATTGGCCATCAAACTGATGTTGATGGTTA 305  
Db 668 GCATGGCGCGCGCAACTGCTGTTCCGCGGATA 698

RESULT 7  
US-09-252-991A-7643/c

; Sequence 7643, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7643  
; LENGTH: 834  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7643

Query Match 7.1%; Score 61.8; DB 4; Length 834;  
Best Local Similarity 54.6%; Pred. No. 9.7e-09;  
Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;  
Qy 35 TGATGGCAATCGTGTGACCTGTGACCCGAATGCGGTATTTGGACGCGTTTGTGTTTA 94  
Db 386 TGGCGGCACTGGCGGTACCTGCTCAACCCCACTCTATCTCGATACCGTACTTCTGA 327  
Qy 95 TCGCGGCGTTCGCGCGCAATACGCGGCAACCGGACGCGTGGATTTTCGCGCGCTGGCGGT 154  
Db 326 TCGGTTCGCTCGCGCGCCAGCAGG-----CGCGCGCGCGCGCGCGCGCGCGCGCG 273  
Qy 155 TCGCGGCAAGCTGATCTGCTTCCGCTGGTGGTTCGCGCGCAGCAGCATTTGTCAAGCC 214  
Db 272 CACGCGCTCGCTGATGTTCTTCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCG 213  
Qy 215 CGCTGTCCAGCCCAAGTGTGGCGCTGATCAACGCTGCGTGGCAGTTGTGATGACCG 274  
Db 212 GCGTGGCG 153  
Qy 275 CATTGGCCATCAAACTGATGTTGATGGTTA 305  
Db 152 GCATGGCGCGCGCAACTGCTGTTCCGCGGATA 122

RESULT 8  
US-09-252-991A-7841  
; Sequence 7841, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7841  
; LENGTH: 894  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7841

Query Match 7.1%; Score 61.8; DB 4; Length 894;  
Best Local Similarity 54.6%; Pred. No. 1e-08;  
Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;  
Qy 35 TGATGGCAATCGTGTGACCTGTGACCCGAATGCGGTATTTGGACGCGTTTGTGTTTA 94

```

Db 339 TGGCGGCACTGGGGTACCTGCTCAACCCCGACGCTATCTCGATACCGTACTTCTGA 398
QY 95 TCGCGGCGCTGGCGCGCAATPAGGCGACACCGAAGGTGATTTTCGGCGCTGGCGGT 154
Db 399 TCGGTTCGCTGGCGCGCCAGCAGG-----CGCGCGCGCGCTATGCGCTCGCGCGCG 452
QY 155 TCGCGGCAAGCTGATCTGTTCCGCTGGTGGGCGCAGCAGCATTTGCAAGCC 214
Db 453 CACGCGCTCGCTGATGTTCTTGCGCTCGCCCTCGCGCGCGCATGGTGGCCCT 512
QY 215 CGCTGTCCAGCCCCAAGGTGGCGCTGGATCAAGCTGCTGGCGAGTTGATGACGG 274
Db 513 GCTGGCGCGCGCGCCACCTGGCGCTGCTGACCTGATGTTGGCGCGCATGATGCTGG 572
QY 275 CATTGGCCATCAACTGATGTTGATGGTTA 305
Db 573 GCATGGCGCGCAACTGCTGTTCCGGGGATA 603

RESULT 9
US-09-489-039A-5370/c
; Sequence 5370, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5370
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5370

Query Match 7.0%; Score 61.2; DB 4; Length 1095;
Best Local Similarity 51.2%; Pred. No. 1.7e-08;
Matches 209; Conservative 0; Mismatches 178; Indels 21; Gaps 2;

QY 469 CTGGTCACGGTATCCGGCGCTACTCTCTGCTCGCGCAGCACCACCCATGCAAGCGCCATC 528
Db 998 CTGGCGCGATCAGCACCAGGTGACTTTCTCTTGGCGCAACACCCAGGCCAGCGCCATC 939
QY 529 TGGCGAAGTGAATGCGCGGCTTCTGGCGCATGCTATTGAGCTTGGGACCATATCAATA 588
Db 938 TGGCAGACCTTCTGCGCCAGCGCTTCCGCTGCTTAAAGCTGACGAATTTTCTTAAT 879
QY 589 TTGTTTCAGTTTCAATGCTCCCTCAGACAGGAGTACTACCTGGCT---GGCGGGGACCC 645
Db 878 CGGGCTGAGTCACTGCTCGCGCTCGGCTGAGAAAGACTGCTGCTGGCGCGCGCGAGTCG 819
QY 646 TCTGGAATTCACATGAGATATTGTCCTGAGCAGCGCCCTGCGCAAGTGTGAGAAAGCA 705
Db 818 GCGGGAATGCGGTTGAGATAGCGTTCGTAAGTTGCGCCCGCGCGCAGCGGTGAAACGCT 759
QY 706 ATGACGCCAAGACATTTGTTGGCAGTCACTGACCAAGTTCTTACGCTATCGCCCGGT 765
Db 758 ATGCTTCCAATCCCTTCGCTTGCAGAAATCCAGCAATCC-----T 717
QY 766 TCTTCCACCCACCATTAATGATGAATAGCTTGGCTGATGAATCAGAAAGCGGCGGCC 825
Db 716 TCTTCCACCGCGGTTGAAACATCGAGTAGCGAGCTGTTGATGATGAGGCGGTGCGG 657
QY 826 TCTTCCGCCATGAATCAGCGCGCTCCGCTGTGAGCTCTGGACCGTAG 873
Db 656 AGGTGCTTAAGGATCTTACCGCGCTCCCGCGCTGGGCCAGAGGATAG 609

RESULT 10

```

```

US-09-894-844-11
; Sequence 11, Application US/09894844
; Patent No. 6886166
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; TITLE OF INVENTION: the M. Tuberculosis Complex
; FILE REFERENCE: STAN1020N
; CURRENT APPLICATION NUMBER: US/09/894,844
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-09-894-844-11

Query Match 6.5%; Score 57; DB 4; Length 597;
Best Local Similarity 53.1%; Pred. No. 2.3e-07;
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 26 AGCCATCTTGATGGCAATCGTGTGACCTGGTGAACCGGATGCTATTGGACGGT 85
Db 326 AGCTCTCTGTGACTGTGGCGCATTCACGTTCTCAACCCACACGCTACTCTGACACGG 385
QY 86 TTGTGTTATTCGGCGCGCTGCGCGCAATACGCGCACACCGACCGGATGATTTGCGCG 145
Db 386 TCGTGTTCAGCGCGCTGCGCAACGAGCACAGCA---CCAGCGCTGGCTGTTGCGCC 442
QY 146 CTGGCGGCTTCGGCGCAAGCCTGATCTGTTCCCGTGGTGGTTCGGCGCAGCAGCAT 205
Db 443 TCGGGCGGCTCACAGCAGTGGCGGTATGTTTCGCCACCCCTCGGGTTCGGAGCGCGCGGT 502
QY 206 TGTTCAGCGCCCTGTCAGCGCCCAAGGTGCGCGCTGGATCAAGCTGCTGTCGCGAGTTG 265
Db 503 TGGCGGGCTGTTTCCACCAACCCCGCTGCTGGAGAAATCTCTGACGCGCTGATCGGCTCA 562
QY 266 TGAATGACCGCATTTGGCCATCAAACTGATGTTGA 298
Db 563 TGAATGCTGGCTGGGAATCTGCTGACCGTGA 595

RESULT 11
US-08-390-878-17/c
; Sequence 17, Application US/08390878
; Patent No. 5700583
; GENERAL INFORMATION:
; APPLICANT: Stover, Charles K.
; APPLICANT: Mahairas, Gregory G.
; TITLE OF INVENTION: VIRULENCE-ATTENUATING GENETIC DELETIONS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower, 20th
; STREET: Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/390,878  
FILING DATE: 17-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunter, Tom  
REGISTRATION NUMBER: 38,498  
REFERENCE/DOCKET NUMBER: 15371A-17  
TELEPHONE: 415/543/9600  
TELEFAX: 415/543/5043  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15239 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-390-878-17

Query Match 6.5%; Score 57; DB 1; Length 15239;  
Best Local Similarity 53.1%; Pred. No. 1.1e-06;  
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;  
QY 26 AGCCCATGTTGATGCAATGCTGCTGACCTGTTGAACCGGAATGCGTATTGACGCGT 85  
Db 4792 AGGTCCTGGTACCTGTGGCGCATTCACGTTCTCAACCCACACGCTACCTCGACACCG 4733  
QY 86 TTGTGTTTATCGCGCGCGTGGCGCGCAATACGGCGACACCGGAGCGTGGATTTCGCGG 145  
Db 4732 TCGTGTGTTAGCGCGCTGGCCACACGACACAGCGA---CCAGCGCTGGCTGTTCGGCC 4676  
QY 146 CTGGCGGCTTCGCGCGAGCTGATCTGTTCCCGCTGGTGGTTCGGCGCGCAGCAT 205  
Db 4675 TCGCGCGCTTCACGACGATGCGGTATGTTCCGCCACCTTCGGGTTTCGGAGCGCGCGT 4616  
QY 206 TGTCAAGCCCGCTGTCAGCGCCCAAGGTGTGGCGCTGGATCAACGTCCTGTCGGCAGTTG 265  
Db 4615 TCGCGCGCTGTTCAACCAACCGCGCTGTGGAGAATCTCGACGCGCTGATCGCGTCA 4556  
QY 266 TGATGACCCATGCGCATCAAACTGATGTGA 298  
Db 4555 TGATGTTGCGTGGGAATCTCGCTGACCGTGA 4523

RESULT 12  
US-09-103-840A-2  
Sequence 2, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 4403765  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
FEATURE:  
OTHER INFORMATION: CDC 1551  
OTHER INFORMATION: "n" bases at various positions throughout the sequence  
OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 6.5%; Score 57; DB 3; Length 4403765;  
Best Local Similarity 53.1%; Pred. No. 1.7e-05;  
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 26 AGCCCATGTTGATGCAATGCTGCTGACCTGTTGAACCGGAATGCGTATTGACGCGT 85  
Db 2227633 AGGTCCTGGTACCTGTGGCGCATTCACGTTCTCAACCCACACGCTACCTCGACACCG 2227692  
QY 86 TTGTGTTTATCGCGCGCGTGGCGCGCAATACGGCGACACCGGAGCGTGGATTTCGCGG 145  
Db 2227693 TCGTGTGTTAGCGCGCTGGCCACACGACACAGCGA---CCAGCGCTGGCTGTTCGGCC 2227749  
QY 146 CTGGCGGCTTCGCGCGAGCTGATCTGTTCCCGCTGGTGGTTCGGCGCGCAGCAT 205  
Db 2227750 TCGCGCGCGTTCACAGCAGTGGGTATGTTCCGCCACCTTCGGGTTTCGGAGCGCGCGT 2227809  
QY 206 TGTCAAGCCCGCTGTCAGCGCCCAAGGTGTGGCGCTGGATCAACGTCCTGTCGGCAGTTG 265  
Db 2227810 TCGCGCGCGCTGTTCAACCAACCGCGCTGTGGAGAATCTCTCGACGCGCTGATCGCGGTCA 2227869  
QY 266 TGATGACCCATGCGCATCAAACTGATGTGA 298  
Db 2227870 TGATGTTGCGTGGGAATCTCGCTGACCGTGA 2227902

RESULT 13  
US-09-103-840A-1  
Sequence 1, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 6.5%; Score 57; DB 3; Length 4411529;  
Best Local Similarity 53.1%; Pred. No. 1.7e-05;  
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;  
QY 26 AGCCCATGTTGATGCAATGCTGCTGACCTGTTGAACCGGAATGCGTATTGACGCGT 85  
Db 2230334 AGGTCCTGGTACCTGTGGCGCATTCACGTTCTCAACCCACACGCTACCTCGACACCG 2230393  
QY 86 TTGTGTTTATCGCGCGCGCTGGCGCGCAATACGGCGACACCGGAGCGTGGATTTCGCGG 145  
Db 2230394 TCGTGTGTTAGCGCGCTGGCCACACGACACAGCGA---CCAGCGCTGGCTGTTCGGCC 2230450  
QY 146 CTGGCGGCTTCGCGCGAGCTGATCTGTTCCCGCTGGTGGTTCGGCGCGCAGCAT 205  
Db 2230451 TCGCGCGCGTTCACAGCAGTGGGTATGTTTCGCCACCTTCGGGTTTCGGAGCGCGCGT 2230510  
QY 206 TGTCAAGCCCGCTGTCAGCGCCCAAGGTGTGGCGCTGGATCAACGTCCTGTCGGCAGTTG 265  
Db 2230511 TCGCGCGCTGTTCAACCAACCGCGCTGTGGAGAATCTCTCGACGCGCTGATCGCGGTCA 2230570  
QY 266 TGATGACCCATGCGCATCAAACTGATGTGA 298  
Db 2230571 TGATGTTGCGTGGGAATCTCGCTGACCGTGA 2230603

RESULT 14  
US-09-252-991A-7571/c  
Sequence 7571, Application US/09252991A



; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7571  
; LENGTH: 1176  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7571

Query Match 6.2%; Score 53.8; DB 4; Length 1176;  
Best Local Similarity 54.3%; Pred. No. 2.9e-06;  
Matches 134; Conservative 0; Mismatches 107; Indels 6; Gaps 1;  
QY 59 TGAACCCGAATGCTATTTGGACGCGTTTGTATTATCGGCGGTGCGCGGCAATACG 118  
DB 1176 TCAACCCCGACGCTATCTCGATACCGTACTTCTGATCGGTTGCTCGGCGCCAGCAGG 1117  
QY 119 GCGACACCGGACGGTGGATTTTCGCGCGTGGCGGTTTCGCGGCAAGCTGATCTGTTCC 178  
DB 1116 -----CGGCGCGGCGGCTATGCCCTCGCGCGCGCCAGCGCCCTCGCTGATGTTCT 1063  
QY 179 CGCTGGTGGGTTTCGCGGACGAGCATTTGTCACGCGCGCTGTCAGCCGCCAAGGTGTCG 238  
DB 1062 TCGCCCTCGCCCTCGGCGGCGCATGGCTGCGCCCTTGGCTGGCGGCGCCGACCTGGC 1003  
QY 239 GCTGGATCAACGTCGTGGGAGTTGTGATGACCGCATTTGCCCATCAAACTGATGTGA 298  
DB 1002 GCCTGCTCGACCTGATGTTGGCGGCGCATGATGCTGGGCATGGCGCGCACTGCTGTTCC 943  
QY 299 TGGGTTA 305  
DB 942 GGGGATA 936

RESULT 15  
US-09-724-623-24/c  
; Sequence 24, Application US/09724623  
; Patent No. 6476209  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Lubbers, Mark W  
; APPLICANT: Dekker, James  
; TITLE OF INVENTION: Polynucleotides, materials incorporating  
; TITLE OF INVENTION: them, and methods for using them.  
; FILE REFERENCE: 1048U1  
; CURRENT APPLICATION NUMBER: US/09/724,623  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 124  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 1002  
; TYPE: DNA  
; ORGANISM: Lactobacillus rhamnosus  
US-09-724-623-24

Query Match 5.8%; Score 50.2; DB 4; Length 1002;  
Best Local Similarity 49.3%; Pred. No. 3.2e-05;  
Matches 176; Conservative 0; Mismatches 163; Indels 18; Gaps 1;  
QY 365 AAATCTCATCGATCGCCTCCAACTCGGCGTCAGAAACTCCAGTTGTTGAGTGAATCAA 424  
DB 994 AAATATCATCAATTTGTTGAATCTCTTCAGCAGTAAGGTGAGATGTTCCGTTGCTTTAA 935

QY 425 GGCTGTGTCGCGTGTCACTGACGAGCAACCAATCAATGCACTGCTCAGGTATCCG 484  
DB 934 GGTATCTCTGAAGGTGTTCACTGAGGTGCTGCCAATGATCAGCTTGTGACAA----- 881  
QY 485 CGCGGTACTCTCTCTGCTCGCGACGACCCCATGCAAGCGCCATCTGCGCAAGTACTGCC 544  
DB 880 -----CGGATCCCGTAAACACCAACCGCCCAAGGCCATTTGACTCAGGGTTTGGT 833  
QY 545 CGCGTTCCTGGCGGATGTCATTGAGCTTGGGACCATATCAATATTCTTCAGGTTCAACA 604  
DB 832 CAGGATCATGCGCAATTTCAATTAAGCGCATTTAGTTGCTTAAACACAGCCCTTTTGGCCT 773  
QY 605 TGCCTCTCAGACAGGGACTTACCCCTGGCTGGCGGGGAAACCTCTTGGAAATTCATCGAGAT 664  
DB 772 TAGCAAAAGTGGCCTTTGTTGGTTGGATGATTTTGAAGTATCCGGAATTCCTTTAGGT 713  
QY 665 ATTTGTCGTGAGCAGGCCCTCGGCAAGTGTGAGAAAGCATGACGCCAAGCCAT 721  
DB 712 AGCGATCTGATACCAAGCCCTTCTGATTAACGCTCCGCTATGCAATCAACCCGCGCACCAT 656

Search completed: March 16, 2004, 03:01:13  
Job time : 82.7645 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 17, 2004, 00:51:03 ; Search time 312.967 Seconds  
10269.549 Million cell updates/sec

Title: US-09-105-117K-1\_COPY\_1421\_2293  
Perfect score: 873  
Sequence: 1 gtcgataagcagcggtttg.....ctgtgagctctggaccgtag 873

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2432557 seqs, 184079884 residues  
Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match %	Length	DB ID	Description
1	873	100.0	2374	14	US-10-196-232-24	Sequence 24, Appl
2	873	100.0	3309400	9	US-09-738-626-1	Sequence 1, Appl
3	568	65.1	627	9	US-09-738-626-3454	Sequence 3454, Appl
4	404.2	46.3	993	12	US-10-627-476-669	Sequence 669, Appl
5	326	37.3	822	10	US-09-746-660A-51	Sequence 51, Appl
6	306	35.1	711	14	US-10-166-142-7	Sequence 7, Appl
7	304.4	34.9	712	14	US-10-166-142-9	Sequence 9, Appl
8	303	34.7	708	9	US-09-738-626-3455	Sequence 3455, Appl
9	253.6	29.0	1095	12	US-10-282-122A-17521	Sequence 17521, A
10	136.6	15.6	1026	15	US-10-369-493-35374	Sequence 35374, A
11	136.6	15.6	1026	15	US-10-369-493-38826	Sequence 38826, A
12	136.6	15.6	1028	15	US-10-369-493-38124	Sequence 38124, A
13	105.8	12.1	1041	12	US-10-282-122A-13437	Sequence 13437, A
14	96.8	11.1	1107	12	US-10-282-122A-14390	Sequence 14390, A
15	96.2	11.0	1041	12	US-10-282-122A-11887	Sequence 11887, A

c 16	90.4	10.4	1026	15	US-10-369-493-37651	Sequence 37651, A
c 17	89.6	10.3	993	14	US-10-156-761-1151	Sequence 1151, Appl
c 18	89.6	10.3	9025608	14	US-10-156-761-1	Sequence 1, Appl
c 19	80.4	9.2	1038	12	US-10-282-122A-19965	Sequence 19965, A
c 20	79.4	9.1	1041	12	US-10-282-122A-7020	Sequence 7020, Appl
c 21	78.2	9.0	1035	12	US-10-282-122A-33379	Sequence 33379, A
c 22	76.4	8.8	999	12	US-10-282-122A-39389	Sequence 39389, A
c 23	74.8	8.6	999	12	US-10-282-122A-38764	Sequence 38764, A
c 24	70.8	8.1	745	15	US-10-260-238-1450	Sequence 1450, Appl
c 25	65.4	7.5	988	9	US-09-974-300-77	Sequence 77, Appl
c 26	64.4	7.4	999	12	US-10-282-122A-12754	Sequence 12754, A
c 27	62.8	7.2	609	14	US-10-156-761-1109	Sequence 1109, Appl
c 28	62.8	7.2	9025608	14	US-10-156-761-1	Sequence 1, Appl
c 29	59.2	6.8	990	12	US-10-282-122A-23370	Sequence 23370, A
c 30	57	6.5	597	9	US-09-894-844-11	Sequence 11, Appl
c 31	57	6.5	597	15	US-10-388-902-11	Sequence 11, Appl
c 32	56.2	6.4	86114	14	US-10-080-170-648	Sequence 648, Appl
c 33	52.4	6.0	990	12	US-10-282-122A-41645	Sequence 41645, A
c 34	50.2	5.8	1002	12	US-10-288-930-24	Sequence 24, Appl
c 35	50.2	5.8	1277	9	US-09-971-536-15	Sequence 15, Appl
c 36	47	5.4	927	15	US-10-369-493-28453	Sequence 28453, A
c 37	47	5.4	927	15	US-10-369-493-31213	Sequence 31213, A
c 38	45.8	5.2	987	12	US-10-282-122A-21647	Sequence 21647, A
c 39	42.2	4.8	969	12	US-10-282-122A-14512	Sequence 14512, A
c 40	41	4.7	8473	16	US-10-038-854-39	Sequence 39, Appl
c 41	41	4.7	8487	16	US-10-038-854-41	Sequence 41, Appl
c 42	41	4.7	8645	16	US-10-038-854-37	Sequence 37, Appl
c 43	41	4.7	8675	16	US-10-038-854-35	Sequence 35, Appl
c 44	40.4	4.6	1392	15	US-10-369-493-31368	Sequence 31368, A
c 45	40.4	4.6	1443	15	US-10-369-493-28610	Sequence 28610, A

ALIGNMENTS

RESULT 1  
US-10-196-232-24  
; Sequence 24, Application US/10196232  
; Publication No. US20030113899A1  
; GENERAL INFORMATION:  
; APPLICANT: YAMAGUCHI, MIKIKO  
; APPLICANT: ITO, HISAO  
; APPLICANT: GUNJI, YOSHIYA  
; APPLICANT: YASUEDA, HISASHI  
; TITLE OF INVENTION: METHOD FOR PRODUCING L-ARGININE  
; FILE REFERENCE: 225391US0  
; CURRENT APPLICATION NUMBER: US/10196,232  
; CURRENT FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: JP 2001-224586  
; PRIOR FILING DATE: 2001-07-25  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 2374  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1025)..(1723)  
; OTHER INFORMATION:  
US-10-196-232-24

Query Match 100.0%; Score 873; DB 14; Length 2374;  
Best Local Similarity 100.0%; Pred. No. 1.2e-272;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GTCGATAGCAGCGGTTTGGTAAAGCCCATGTTGATGCGCAATCGTGTGACCTGTTG 60  
Db 1421 GTCGATAGCAGCGGTTTGGTAAAGCCCATGTTGATGCGCAATCGTGTGACCTGTTG 1480  
Qy 61 AACCCGAAATCGTATTTCGACGCGTTTGTGTTTATTCGCGCGCGTTCGCGCGCAATACGCG 120  
Db 1481 AACCCGAAATCGTATTTCGACGCGTTTGTGTTTATTCGCGCGCGTTCGCGCGCAATACGCG 1540



QY 841 TCAGCGGCTCGGCTGTGAGCTCTGACCGTAG 873  
 Db 1327708 TCAGCGGCTCGGCTGTGAGCTCTGACCGTAG 1327676

RESULT 3

US-09-738-626-3454/c  
 ; Sequence 3454, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIALI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAOKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; PRIOR FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: Patent in ver. 3.0  
 ; SEQ ID NO 3454  
 ; LENGTH: 627  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-3454

Query Match 65.1%; Score 568; DB 9; Length 627;

Best Local Similarity 100.0%; Pred. No. 8.3e-174;  
 Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 GTTTTCGGGGTTTGGAAATCGGTGCGCTTCGCCCAATGTTGATCGCGGCTGTGGGA 365  
 Db 627 GTTTTCGGGGTTTGGAAATCGGTGCGCTTCGCCCAATGTTGATCGCGGCTGTGGGA 568  
 QY 366 AATCTCATCGATCGCGCTCCAACTCGCGGTGAGAAACTCCAAAGTTTGTGAGTGAATCAAG 425  
 Db 567 AATCTCATCGATCGCGCTCCAACTCGCGGTGAGAAACTCCAAAGTTTGTGAGTGAATCAAG 508  
 QY 426 GCTGTTGTCAGCTGCTCAACTGCAAGACCAATCAATGCACTGCTCAGCGTATCCGC 485  
 Db 507 GCTGTTGTCAGCTGCTCAACTGCAAGACCAATCAATGCACTGCTCAGCGTATCCGC 448  
 QY 486 GCGTACTCTCTCTGCTCGGCGAGCACCCATGCGGCGCATCTCGCAAGTCACTGCGCC 545  
 Db 447 GCGTACTCTCTCTGCTCGGCGAGCACCCATGCGGCGCATCTCGCAAGTCACTGCGCC 388  
 QY 546 GCGTCTCTGCGGCGATGTCATGAGCTTGGGACCATATCAATATTGTTTCACTTCAACAT 605  
 Db 387 GCGTCTCTGCGGCGATGTCATGAGCTTGGGACCATATCAATATTGTTTCACTTCAACAT 328  
 QY 606 GCGTCTGAGACGAGGACTTACCTGCTGCGGCGGACCTCTGGAAATTCATCGAGATA 665  
 Db 327 GCGTCTGAGACGAGGACTTACCTGCTGCGGCGGACCTCTGGAAATTCATCGAGATA 268  
 QY 666 TTTTCTCGGTGAGACGCGCTGCGCAAGTGTGAGAAAGCAATGAGCCCAAGACCATTTGTT 725  
 Db 267 TTTTCTCGGTGAGACGCGCTGCGCAAGTGTGAGAAAGCAATGAGCCCAAGACCATTTGTT 208  
 QY 726 GSCAGCTGATGCAACAAGTTTCTCAGCGTCAATCGCGGTTCTCTCACCCCAAGCAATTAAT 785

Db 207 GGCAGCTGACTGCAACAAGTTTCTCACCCTCATCGCGGTTCTCTCCACCAACGAAATTAAT 148  
 QY 786 GATGGAATAGCTTGCGTGTGATCAATCAGAAGCGGCGGCGGCTCTCGGCATGAACTCAGC 845  
 Db 147 GATGGAATAGCTTGCGTGTGATCAATCAGAAGCGGCGGCGGCTCTCTCGGCATGAACTCAGC 88  
 QY 846 GCGCTCGGCTGTGAGCTCTCGGACCGTAG 873  
 Db 87 GCGCTCGGCTGTGAGCTCTCGGACCGTAG 60

RESULT 4

US-10-627-476-669/c  
 ; Sequence 669, Application US/10627476  
 ; Publication No. US20040030116A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pompejus, Mark  
 ; APPLICANT: Kroger, Burkhard  
 ; APPLICANT: Schoder, Hartwig  
 ; APPLICANT: Zelder, Oskar  
 ; APPLICANT: Haberhauer, Gregor  
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
 ; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE  
 ; FILE REFERENCE: BGI-125CPCN  
 ; CURRENT APPLICATION NUMBER: US/10/627,476  
 ; CURRENT FILING DATE: 2003-07-25  
 ; PRIOR APPLICATION NUMBER: 09/602,787  
 ; PRIOR FILING DATE: 2000-06-23  
 ; PRIOR APPLICATION NUMBER: USN 60/141031  
 ; PRIOR FILING DATE: 1999-06-25  
 ; PRIOR APPLICATION NUMBER: DE 19931454.3  
 ; PRIOR FILING DATE: 1999-07-08  
 ; PRIOR APPLICATION NUMBER: DE 19931478.0  
 ; PRIOR FILING DATE: 1999-07-08  
 ; PRIOR APPLICATION NUMBER: DE 19931563.9  
 ; PRIOR FILING DATE: 1999-07-08  
 ; PRIOR APPLICATION NUMBER: DE 19932122.1  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932124.8  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932125.6  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932128.0  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932180.9  
 ; PRIOR FILING DATE: 1999-07-09  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 678  
 ; SEQ ID NO 669  
 ; LENGTH: 993  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (101)..(970)  
 ; OTHER INFORMATION: RXN03164  
 US-10-627-476-669

Query Match 46.3%; Score 404.2; DB 12; Length 993;

Best Local Similarity 98.8%; Pred. No. 1.7e-120;  
 Matches 418; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 453 AGCACCAATCAATGACACTGGTCAAGGTA--TCGCGCGGTACTCTCTTGTCTCGCGCAGC 510  
 Db 993 AGCACCAATCAATGACACTGGTCAAGGTAATCGCGCGGTACTCTCTTGTCTCGCGCAGC 934  
 QY 511 ACCCATCAAGCGCCATCTCGCAAGTGAATCGCGCGGTTCTCGGCGGATGTCATTGAGC 570  
 Db 933 ACCCATCAAGCGCCATCTCGCAAGTGAATCGCGCGGTTCTCGGCGGATGTCATTGAGC 874  
 QY 571 TTGCGGACCAATCAATGTCATTGTCACGTTCAACATGCGCTCAGACGGGACTTACCTGCG 630

Db 873 TTGGGACCATATTAATTTTCAAGTTCACATGCTCCCTCAGACAGGACTTACCCCTGG 814  
 QY 631 CTGGCGGGGAACTCTTGGATTCATCGAGATATTTGCTGAGAGGCCCTTGGCA 690  
 Db 813 CTGGCGGGGAACTCTTGGATTCATCGAGATATTTGCTGAGAGGCCCTTGGCA 754  
 QY 691 AGTGTGTGAGAAAGCAATGACGCCAAGACCAATTTGTCGAGCTGACTGCAACAAAGTTCTCA 750  
 Db 753 AGTGTGTGAGAAAGCAATGACGCCAAGACCAATTTGTCGAGCTGACTGCAACAAAGTTCTCA 694  
 QY 751 CCGTCATCGCCCGGCTCTCCACCCCAAGCAATTAATGATGGAATAGCTGGCTGATGAATC 810  
 Db 693 CCGTCATCGCCCGGCTCTCCACCCCAAGCAATTAATGATGGAATAGCTGGCTGATGAATC 694  
 QY 811 AGAAGCGGGCAGCCCTCTCCGCCCATGAATCTCAGCGGCTCCGCTGTGAGCTCTGGACCG 870  
 Db 633 AGAAGCGGGCAGCCCTCTCCGCCCATGAATCTCAGCGGCTCCGCTGTGAGCTCTGGACCG 574  
 QY 871 TAG 873  
 Db 573 TAG 571

RESULT 5  
 US-09-746-660A-51  
 ; Sequence 51, Application US/09746660A  
 ; Publication No. US20030049804A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pompeius, Markus  
 ; APPLICANT: Kroger, Burkhard  
 ; APPLICANT: Schroder, Hartwig  
 ; APPLICANT: Zelder, Oskar  
 ; APPLICANT: Haberhauer, Gregor  
 ; APPLICANT: Kim, Jun-Won  
 ; APPLICANT: Lee, Heung-Schick  
 ; APPLICANT: Hwang, Byung-Joon  
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING  
 ; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS  
 ; FILE REFERENCE: BG1-121CP2  
 ; CURRENT APPLICATION NUMBER: US/09/746,660A  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 09/606740  
 ; PRIOR FILING DATE: 2000-06-23  
 ; PRIOR APPLICATION NUMBER: 09/603124  
 ; PRIOR FILING DATE: 2000-06-23  
 ; PRIOR APPLICATION NUMBER: 60/141031  
 ; PRIOR FILING DATE: 1999-06-25  
 ; PRIOR APPLICATION NUMBER: 60/142101  
 ; PRIOR FILING DATE: 1999-07-02  
 ; PRIOR APPLICATION NUMBER: 60/148613  
 ; PRIOR FILING DATE: 1999-08-12  
 ; PRIOR APPLICATION NUMBER: 60/187970  
 ; PRIOR FILING DATE: 2000-03-09  
 ; PRIOR APPLICATION NUMBER: DE 19931420.9  
 ; PRIOR FILING DATE: 1999-07-08  
 ; NUMBER OF SEQ ID NOS: 125  
 ; SOFTWARE: PatentIn Vers. 2.0  
 ; SEQ ID NO 51  
 ; LENGTH: 822  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (101)..(799)  
 ; OTHER INFORMATION: RXA01394  
 US-09-746-660A-51  
 Query Match 37.3%; Score 326; DB 10; Length 822;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-95;  
 Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GTCGATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGCTGACCTGGTTG 60

Db 497 GTCGATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGCTGACCTGGTTG 556  
 QY 61 AACCCGAATCGGTATTTTGGACCGCTTTGTTTATCGCGGCGTCCGCGCAATACGCGC 120  
 Db 557 AACCCGAATCGGTATTTTGGACCGCTTTGTTTATCGCGGCGTCCGCGCAATACGCGC 616  
 QY 121 GACACCGGACGCGTGGATTTTTCGCCGCTTGGCGCGTTCGCGGCAAGCCTGATCTGTTCCCG 180  
 Db 617 GACACCGGACGCGTGGATTTTTCGCCGCTTGGCGCGTTCGCGGCAAGCCTGATCTGTTCCCG 676  
 QY 181 CTGGTGGGTTTCGCCGCGACGACGATTTGACGCCGCTGTCCAGCCCAAGGTGTGGCGC 240  
 Db 677 CTGGTGGGTTTCGCCGCGACGACGATTTGACGCCGCTGTCCAGCCCAAGGTGTGGCGC 736  
 QY 241 TGGATCAACGTCGTCGTTGGCAGTTCGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300  
 Db 737 TGGATCAACGTCGTCGTTGGCAGTTCGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 796  
 QY 301 GGTAGTTTTCGCCGCGGTTTGGAAATC 326  
 Db 797 GGTAGTTTTCGCCGCGGTTTGGAAATC 822

RESULT 6  
 US-10-166-142-7  
 ; Sequence 7, Application US/10166142  
 ; Publication No. US20030124687A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GUNJI, YOSHIYA  
 ; APPLICANT: YASUEDA, HISASHI  
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL  
 ; TITLE OF INVENTION: ASSIMILATING BACTERIUM  
 ; FILE REFERENCE: 223789US  
 ; CURRENT APPLICATION NUMBER: US/10/166,142  
 ; CURRENT FILING DATE: 2002-06-11  
 ; PRIOR APPLICATION NUMBER: JP 2001-1777075  
 ; PRIOR FILING DATE: 2001-06-12  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 7  
 ; LENGTH: 711  
 ; TYPE: DNA  
 ; ORGANISM: Brevibacterium lactofermentum  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(711)  
 ; OTHER INFORMATION:  
 US-10-166-142-7  
 Query Match 35.1%; Score 306; DB 14; Length 711;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-88;  
 Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GTCGATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGCTGACCTGGTTG 60  
 Db 406 GTCGATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGCTGACCTGGTTG 465  
 QY 61 AACCCGAATCGGTATTTTGGACCGCTTTGTTTATCGCGGCGTCCGCGCAATACGCGC 120  
 Db 466 AACCCGAATCGGTATTTTGGACCGCTTTGTTTATCGCGGCGTCCGCGCAATACGCGC 525  
 QY 121 GACACCGGACGCGTGGATTTTTCGCCGCTTGGCGCGTTCGCGCAAGCCTGATCTGTTCCCG 180  
 Db 526 GACACCGGACGCGTGGATTTTTCGCCGCTTGGCGCGTTCGCGCAAGCCTGATCTGTTCCCG 585  
 QY 181 CTGGTGGGTTTCGCCGCGACGACGATTTGACGCCGCTGTCCAGCCCAAGGTGTGGCGC 240  
 Db 586 CTGGTGGGTTTCGCCGCGACGACGATTTGACGCCGCTGTCCAGCCCAAGGTGTGGCGC 645  
 QY 241 TGGATCAACGTCGTCGTTGGCAGTTCGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300  
 Db 646 TGGATCAACGTCGTCGTTGGCAGTTCGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 705

QY 301 GGTAG 306  
Db 706 GGTAG 711

## RESULT 7

US-10-166-142-9  
; Sequence 9, Application US/10166142  
; Publication No. US20030124687A1  
; GENERAL INFORMATION:  
; APPLICANT: GUNJI, YOSHIYA  
; APPLICANT: YASUEDA, HISASHI  
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL  
; FILE REFERENCE: 223789US  
; CURRENT APPLICATION NUMBER: US/10/166,142  
; CURRENT FILING DATE: 2002-06-11  
; PRIOR APPLICATION NUMBER: JP 2001-1777075  
; PRIOR FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 712  
; TYPE: DNA  
; ORGANISM: Brevibacterium lactofermentum  
; NAME/KEY: CDS  
; LOCATION: (1)..(375)  
; OTHER INFORMATION:  
US-10-166-142-9

Query Match 34.9%; Score 304.4; DB 14; Length 712;

Best Local Similarity 99.7%; Pred. No. 3.8e-88;  
Matches 305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTCGATAAGCAGCGGTTTGGGTAAGCCCAATGTTGATGGCAATCGTCTGACCTGGTTG 60  
Db 407 GTCGATAAGCAGCGGTTTGGGTAAGCCCAATGTTGATGGCAATCGTCTGACCTGGTTG 466  
QY 61 AACCCGAATGCGTATTTGGACGCGTTTGTGTTATCGCGCGTGGCGGCGCAATACGGC 120  
Db 467 ARCCCGAATGCGTATTTGGACGCGTTTGTGTTATCGCGCGTGGCGGCGCAATACGGC 526  
QY 121 GACACCGAGCGTGATTTTCGCCGCTGGCGGCTTCGCGCAAGCTGATCTGGTTCCCG 180  
Db 527 GACACCGAGCGTGATTTTCGCCGCTGGCGGCTTCGCGCAAGCTGATCTGGTTCCCG 586  
QY 181 CTGGTGGGTTTCGCGCGCAGCAGCATTTGTCACGCCCGCTGTCCAGCCCCCAAGGTGTGGCGC 240  
Db 587 CTGGTGGGTTTCGCGCGCAGCAGCATTTGTCACGCCCGCTGTCCAGCCCCCAAGGTGTGGCGC 646  
QY 241 TGGATCAAGCTCGTGGCGAGTTGTGATGACCCGATCGCCATCAAACTGATGTTGATG 300  
Db 647 TGGATCAAGCTCGTGGCGAGTTGTGATGACCCGATCGCCATCAAACTGATGTTGATG 706  
QY 301 GGTAG 306  
Db 707 GGTAG 712

## RESULT 8

US-09-738-626-3455  
; Sequence 3455, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOHI, AKIHIRO

; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 3455  
; LENGTH: 708  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3455

Query Match 34.7%; Score 303; DB 9; Length 708;

Best Local Similarity 100.0%; Pred. No. 1.1e-87;  
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGATAAGCAGCGGTTTGGGTAAGCCCAATGTTGATGGCAATCGTCTGACCTGGTTG 60  
Db 406 GTCGATAAGCAGCGGTTTGGGTAAGCCCAATGTTGATGGCAATCGTCTGACCTGGTTG 465  
QY 61 AACCCGAATGCGTATTTGGACGCGTTTGTGTTATCGCGCGTGGCGGCGCAATACGGC 120  
Db 466 AACCCGAATGCGTATTTGGACGCGTTTGTGTTATCGCGCGTGGCGGCGCAATACGGC 525  
QY 121 GACACCGAGCGTGATTTTCGCCGCTGGCGGCTTCGCGCAAGCTGATCTGGTTCCCG 180  
Db 526 GACACCGAGCGTGATTTTCGCCGCTGGCGGCTTCGCGCAAGCTGATCTGGTTCCCG 585  
QY 181 CTGGTGGGTTTCGCGCGCAGCAGCATTTGTCACGCCCGCTGTCCAGCCCCCAAGGTGTGGCGC 240  
Db 586 CTGGTGGGTTTCGCGCGCAGCAGCATTTGTCACGCCCGCTGTCCAGCCCCCAAGGTGTGGCGC 645  
QY 241 TGGATCAAGCTCGTGGCGAGTTGTGATGACCCGATCGCCATCAAACTGATGTTGATG 300  
Db 646 TGGATCAAGCTCGTGGCGAGTTGTGATGACCCGATCGCCATCAAACTGATGTTGATG 705  
QY 301 GGT 303  
Db 706 GGT 708

## RESULT 9

US-10-282-122A-17521/c  
; Sequence 17521, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848

```
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17521
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-17521

Query Match      29.0%; Score 253.6; DB 12; Length 1095;
Best Local Similarity 65.6%; Pred. No. 1.5e-71;
Matches 370; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

QY 310 TCGCGGGTGTGGAAATCGGTGCGCCCAATGTTGATCGCGCGGTGCGGGAAATC 369
DB 1094 TCGTGAACCTTCAGGCGAGTAGACACCGGCCAAATTAATTCACCGCTCTTGGCAACA 1035

QY 370 TCATCGATCGCCTCCAACTCGGCGTCAGAAAATCCAAAGTTGTGAGTGAATCAAGCTG 429
DB 1034 GTGTCAATATAATTTGGTTCTCTCAACGCTGAACCTCAAGATTAATTAAGAGCTCCCAAGATTT 975

QY 430 TTGTCCAGCTGCTCAACTGAGGAGCAGCAACCAATCAATGACTGTGACGGTATCCGCGCGG 489
DB 974 TCATCCAACTGCGCTACCGAGATGCTCCGATCAACGCACTGGTTACCGTTGTGGCCCCA 915

QY 490 TACTCTCTCTGCTGCGGAGCAGCACCATGCAAGCGCATCTGCGCAAGTCACTTGCCCGCGT 549
DB 914 TAAATGCTGTTGTTCCGCGAGCACCAACGCAATAGCCAGTGTGCGCAACGTTTGGCCAGT 855

QY 550 TCCTGGGCGATGTCATGAGCTTGGGACCAATATCAATATTTGTTCAAGTTCAAGTTCAAGTCC 609
DB 854 CGTAAACGCTAAATCATTCAGTGCAGCAGCAGCCATGTCAGATTTTGTAGCGTTGAGCATCTCT 795

QY 610 TCAGACAGGAGCTTACCTGGCTGGCGGGAGACCTCTGGAATTCATCGAGATATTG 669
DB 794 TTGCCCAAGGAGCTTCCCTGCTGCTGCGCGTGAATCAATGCGACGCTCGAGATACTTG 735

QY 670 TCGGTGACAGGCGCTGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCATTTGTGGCA 729
DB 734 TCAGTTAAAGCCCTGTCAGTGGGGAATGCAATACCAACGCTCGTTATTCGCC 675

QY 730 GCTGACTGCAACAGTTCTCACCGTATCCCGGTTCTCCACCAACGATTAATGATG 789
DB 674 CGAGACTCTAACAGTTCTCGCCATCTTCACTGTTCTTACCCAGCGGTTCAAGATA 615

QY 790 GAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCGGCATGAACCTCAGCCGCC 849
DB 614 GAATAGCTCGGTTGATGAATCAGAGCGGCGAGCCCTCACCAGCTAGATTTCCAGATC 555

QY 850 TCCGCTGTGAGCTCTGACCGGTAG 873
DB 554 TCTCGGTTAGTTCAGGTCCTCATAG 531

RESULT 10
US-10-369-493-35374/c

; Sequence 35374, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35374
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-35374

Query Match      15.6%; Score 136.6; DB 15; Length 1026;
Best Local Similarity 56.8%; Pred. No. 1.4e-33;
Matches 312; Conservative 0; Mismatches 204; Indels 33; Gaps 2;

QY 325 TCGGTGGCTTCGCCCAAAATGTTGATCGCGCGTCTGCGGAAATCTCATGATCGCCTCC 384
DB 1022 TCGGCAATTTTCCCGAGGGTTCATATCGCATCTTGGCGTAAACGGTCTGATTTCGGCC 963

QY 385 AACTCGGCGTCAGAAAATCCAAAGTTGTGAGTGAATCAAGGCTGTGTCAGCTGCTCA 444
DB 962 AGCTCTCGGTAGAGAACTCGGCATATCGAGTGTCTTTCACGCACTCTCGACCTGTTTCG 903

QY 445 ACTGACGAGCAGCAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504
DB 902 ACAGGCTTGGCGCAATCAATGCTGAGTGAATTCGCGCGC-----CG 861

QY 505 CGCAGCAGCATCAAGCGCCATCTGCGCAAGTGAATCGCCGCGTTCTCGGCGGAGTGA 564
DB 860 CGCAGAACCCAGGCAATTCGCATCTGCGCGAGCGTCTGCGCAGCGCTCGGCAATGCTG 801

QY 565 TTGAGCTTGGGCGCAATATCAATATTTGTTCAATGTTTCAATGCTGCTGCTGCTGCTGCTG 624
DB 800 TTTCAGCGCGCGGATGTTTTTCGACATTCGCTGTTGAGGAGGCGCGGTTGAGTGA 741

QY 625 CCCTGGCTGCGCGGCAACCTCTCTGGAATTCATGAGATATTTGTCGTTGAGCAGGCGCC 684
DB 740 CTCTGTGAGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 681

QY 685 TCGCAAGTGTGAGAAAGCAATGACCGCAAGACCAATTTGTTGGCAGCTGATGCAACAG 744
DB 680 TCGCGCAGCGGCAAGCAATGAGCGCATAC-----CAGTCTTCC 636

QY 745 TTCTCACCGTCACTGCGCGGTTCTTCCACCAACGATTAATGATGGAATGCTTGGCTGA 804
DB 635 AGCGTATCGAAGACCGCTTCTTCTCGATCCAGCGGTTGATCATGGAATGCTCGGCTGG 576

QY 805 TGAATCAGAGCGGCGAGCCCTCTCTCGGCATGAATCACTCAGCGCGCTCGCTGTGAGCTCT 864
DB 575 TGAATGAGGCGGCTGCGGATCTTTCAGGATAGCGGCGCTCGCGGCTCGCGGCTTC 516

QY 865 GCAGCGTAG 873
DB 515 GAGTTGTAG 507

RESULT 11
US-10-369-493-38826/c
; Sequence 38826, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
```



; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; PRIOR FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; LENGTH: 1026  
 ; TYPE: DNA  
 ; ORGANISM: Agrobacterium tumefaciens  
 ; US-10-369-493-38826

Query Match 15.6%; Score 136.6; DB 15; Length 1026;  
 Best Local Similarity 56.8%; Pred. No. 1.4e-33;  
 Matches 312; Conservative 0; Mismatches 204; Indels 33; Gaps 2;  
 QY 325 TCGTGGCCCTTCGCCCAATGTTGATCCCGGCTGTGGGAAATCTCATCGATCGCTCC 384  
 Db 1022 TCGGCAGATTTTGGCCAGAGTTGATATCCGCATCTTGGCGTAAACGGTCGATTCGGCC 963  
 QY 385 AACTCGGCGTCAGAAACTCCCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGCTGTCTCA 444  
 Db 962 AGCTCTCGGTAGAGACTTCGGCATTTATCGAGTGTTCACGAGTCTTCGACCTGTTCG 903  
 QY 445 ACTGAGAGCAACCAATCAATGCACTGGTACAGGTATCCGCGCGTACTCTCTCTTGTCTCG 504  
 Db 902 ACACGGCTTGGCCCAATCAATGCTGAGGTAATTCGGCGCGC-----CG 861  
 QY 505 CGCAGACCCATGCAAGCGCCATCTGGCAAGTGAATGCTGCGCGGTTCCTGGGCGATGTCA 564  
 Db 860 CGCAGAACCCAGGCAATTTGCCATCTCGCCAGCGTCTGCGCCACGCGCTCGGCAATGCTG 801  
 QY 565 TTGAGCTTGGGACCATATCAATATTGTTACAGTTTCAACATGCGCTTCAGACAGGACTTA 624  
 Db 800 TTCAGCGCGGAGATGTTTTCGACATTTGCGCTGTTGAGGAGCGCGGTTGAGTACTTG 741  
 QY 625 CCCTGGTGGCGCGGAAACCTCTGGAATTCATCGAGATATTGTTCCGTGAGCAGGCGCC 684  
 Db 740 CTCTGTGAGGCAAGCTCCCATCCGACACCGCCAGATATTTCGTGTCAGCATGCC 681  
 QY 685 TGGCAAGTGGTGAAGCAATGACCCCAAGACCATTTGTTGCACTGATCTGCAACAG 744  
 Db 680 TGGCCAGCGCGGCAAGACGATGGAGCGGATACC-----CAGTTCTTCC 636  
 QY 745 TTCTCACCGTTCATCGCCGGTTCTCCACCCAACTAATGATGGAATAGCTTGGCTGA 804  
 Db 635 AGCGTATCGACAGACCGTCTTCTCGATCCAGCGTTGATCATCGATAGCTCGGCTGG 576  
 QY 805 TGAATCAGAGCGGCGAGCCCTCTCCGCAATGAACTCAGCGCTCCGCTCGTGAAGTCT 864  
 Db 575 TGGATGATGAGGCGGTGCGAGATCTTCAGGATAGCGGCGCTCGCGGCTCGGCTTC 516  
 QY 865 GGACCGTAG 873  
 Db 515 GAGTTGTAG 507

RESULT 12  
 US-10-369-493-38124/c  
 ; Sequence 38124, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; PRIOR FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; LENGTH: 1028  
 ; TYPE: DNA  
 ; ORGANISM: Agrobacterium tumefaciens  
 ; US-10-369-493-38124

Query Match 15.6%; Score 136.6; DB 15; Length 1028;  
 Best Local Similarity 56.8%; Pred. No. 1.5e-33;  
 Matches 312; Conservative 0; Mismatches 204; Indels 33; Gaps 2;  
 QY 325 TCGTGGCCCTTCGCCCAATGTTGATCCCGGCTGTGGGAAATCTCATCGATCGCTCC 384  
 Db 1021 TCGGCAGATTTTGGCCAGAGTTGATATCCGCATCTTGGCGTAAACGGTCGATTCGGCC 962  
 QY 385 AACTCGGCGTCAGAAACTCCCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGCTGTCTCA 444  
 Db 961 AGCTCTCGGTAGAGACTTCGGCATTTATCGAGTGTTCACGAGTCTTCGACCTGTTCG 902  
 QY 445 ACTGAGAGCAACCAATCAATGCACTGGTACAGGTATCCGCGCGTACTCTCTCTTGTCTCG 504  
 Db 901 ACACGGCTTGGCCCAATCAATGCTGAGGTAATTCGGCGCGC-----CG 860  
 QY 505 CGCAGACCCATGCAAGCGCCATCTGGCAAGTGAATGCTGCGCGGTTCCTGGGCGATGTCA 564  
 Db 859 CGCAGAACCCAGGCAATTTGCCATCTCGCCAGCGTCTGCGCCACGCGCTCGGCAATGCTG 800  
 QY 565 TTGAGCTTGGGACCATATCAATATTGTTACAGTTTCAACATGCGCTTCAGACAGGACTTA 624  
 Db 799 TTCAGCGCGGAGATGTTTTCGACATTTGCGCTGTTGAGGAGCGCGGTTGAGTACTTG 740  
 QY 625 CCCTGGTGGCGCGGAAACCTCTGGAATTCATCGAGATATTGTTCCGTGAGCAGGCGCC 684  
 Db 739 CTCTGTGAGGCAAGCTCCCATCCGACACCGCCAGATATTTCGTGTCAGCATGCC 680  
 QY 685 TGGCAAGTGGTGAAGCAATGACCCCAAGACCATTTGTTGCACTGATCTGCAACAG 744  
 Db 679 TGGCCAGCGCGGCAAGACGATGGAGCGGATACC-----CAGTTCTTCC 635  
 QY 745 TTCTCACCGTTCATCGCCGGTTCTCCACCCAACTAATGATGGAATAGCTTGGCTGA 804  
 Db 634 AGCGTATCGACAGACCGTCTTCTCGATCCAGCGTTGATCATCGATAGCTCGGCTGG 575  
 QY 805 TGAATCAGAGCGGCGAGCCCTCTCCGCAATGAACTCAGCGCTCCGCTCGTGAAGTCT 864  
 Db 574 TGGATGATGAGGCGGTGCGAGATCTTCAGGATAGCGGCGCTCGCGGCTCGGCTTC 515  
 QY 865 GGACCGTAG 873  
 Db 514 GAGTTGTAG 506

RESULT 13  
 US-10-282-122A-13437/c  
 ; Sequence 13437, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant

```

; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13437
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Burkholderia fungorum
US-10-282-122A-13437

Query Match      12.1%; Score 105.8; DB 12; Length 1041;
Best Local Similarity 55.6%; Pred. No. 1.5e-23;
Matches 293; Conservative 0; Mismatches 192; Indels 42; Gaps 3;

QY 331 GCCTTCGCCCAATGTTGATGCGCGCTGCTGGGAATCTCATCGATCGCTCCCACTCG 390
DB 1022 GGCCTTTTCCACAGGTTGATGCCGCTTCCGTCGCGTGTAGCGGTGCGATCTCGGCCAGCTCG 963

QY 391 GCGTCAGAACTCCAAAGTTGTTGAGTGAATCAAGCTGTTGTCACAGCTGCTCAACTGAC 450
DB 962 TCCCTTCGAGAACCGAAGTTCTTCAGCGCGCCGACGTTTTCACGACCTGCTCCGCGCG 903

QY 451 GAAGCACCACCAATCAATGCACTGGTCAACGCTATCCGCGCCGTAATCTCTCTTGTCTCGCGCAGC 510
DB 902 CTCGCGCGGATCAGCACGGAAGTCACG-----CGCGGATCGCGCAGC 861

QY 511 ACCCATGACGCGGATCTGCGCAAGTGAATGCGCGCTTCTCGGCGGATGTCATTGAGC 570
DB 860 GCCCAGCGAGCGCGCATTTGTGCGAGGCTCTGACGCGGACGTTGCGGATGCTGTTGAGC 801

QY 571 TTGCGGACCATATCAATATTGTTTCAAGTTTCAATGCTTCAACATGCTTCAAGCAGGACTTACC----- 626
DB 800 TTGCGCAGCTGCTCGATATTCTGCGGCTCAGATGCTCTGTTCAACGAACCGCGCGCC 741

QY 627 --CTGGCTGGCGCGGGAACCCCTCTGGAATTCATCGAGATATTTGCTCGTGAAGAGGCC 684
DB 740 GGCCTTTGTTGACGCGCGCTTTCCGCGACGCGCTTTCAGATATTGCTCTGTTGAGCAGCCCC 681

QY 685 TGGCGAGTGTGAGAACGATGACGCAAGACCATTTGTTGGCAGCTGACTGCAACAAAG 744
DB 680 TGAGCAGCGCGGTGAACGCGGATGCGCGCGCGCGCGCGCTTCTCAATGCGCGCGCAAA-- 622
QY 745 TTCTCACCGTCATCGCCCGGCTTCTTCCACCAAGTAAATGATGGAATAGTCTGGCTGA 804
DB 621 -----CTCATGCTCGATCCAGCATGTTGTAAGAGGCTGA 579

QY 805 TGAATCAGAACGCGGCGAGCCCTCTCTCCGCCATGAATCAAGCCGCTC 851

```

```

DB 578 TGAATCAGACGCGGACCTTGTATTCCGGGACGACCTTCGCCATTTC 532
|||||
RESULT 14
US-10-282-122A-14390/c
; Sequence 14390, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Chlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14390
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Burkholderia mallei
US-10-282-122A-14390

Query Match      11.1%; Score 96.8; DB 12; Length 1107;
Best Local Similarity 56.9%; Pred. No. 1.3e-20;
Matches 232; Conservative 0; Mismatches 152; Indels 24; Gaps 2;

QY 327 GGTGGCTTCGCCCAATGTTGATGCGCGCTGCTGGGAATCTTCATCGATCGCTCCAA 386
DB 1092 GGAGCGCTTTCCCAACAGATTGATCCGCTTCGGTCGCATAGCGTGCATCTCGCGCAT 1033

QY 387 CTCGGCGTCAGAAAATCCCAAGTTGTTGAGTGAATCAAGCTGTTGTCAGCTGCTCAAC 446
DB 1032 CTCGTGCTGCTGAACCGAGATGCGCGCGCGCGAGCTTCTCGCGCACCTGCTCCGC 973

QY 447 TGACGAAGCAACCAATCAATGTCACGCTGTCACGCTATCCGCGCGTACTCTCTTGTCTCGCG 506
DB 972 GCGGCTGCGACCGCATCAACGCGGAGTCAAG-----CGCGAATCGCG 931

QY 507 CAGCACCACCAAGCGCCATCTCGCGCAAGTGAATCGCCGCGGTTCTCTGGCGGATGTCATT 566
DB 930 CAGCACCACCAAGCGCCATCTCGCGCGAGGCTCTGCGCGCGCGCTGCGCGATCTCGTT 871

```

QY 567 GAGCTGGCGGACCATATCAATATTGTTACGTTTCAACATGCGCTCAGACAGGACTTACC 626  
 DB 870 CAGCTGGCGGACCATATCAATATTGTTACGTTTCAACATGCGCTCAGACAGGACTTACC 811  
 QY 627 -----CTGGCTGGCGGCGGAAACCTCTCTGGAATTCATCGAGATATTTGTCGCGAGCAG 680  
 DB 810 GCCCGGCTGTTGATCGCGCATCCGCGGCAACGCGTTTCAGATATCTTCAGAGGTCAGAAAG 751  
 QY 681 GCCCTGGCAAGTGTGGAAGCAATGACGCCAGACCAATTTGTCG 728  
 DB 750 CCCTCTGGCGAGCGCGGTGAACGCGATGACGCCGCGCTGCTGTC 703

RESULT 15

US-10-282-122A-11887/c  
 ; Sequence 11887, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangou  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 11887  
 ; LENGTH: 1041  
 ; TYPE: DNA  
 ; ORGANISM: Burkholderia cepacia  
 US-10-282-122A-11887

Query Match 11.0%; Score 96.2; DB 12; Length 1041;  
 Best Local Similarity 55.7%; Pred. No. 1.9e-20;  
 Matches 241; Conservative 0; Mismatches 168; Indels 24; Gaps 2;  
 QY 327 GGTGGCTTTCGCCAAATGTTGATGCGCGGTGTTGGGAAATCTCATCGATCGCTCCAA 386  
 DB 1026 GGACGGCTTTTCCACAGATTGATGCGCGCTTCCGTCGCGTAAGATCGATCTCCGGAG 967  
 QY 387 CTCGGCGTCAGAAACTCCAAAGTTGTTGAGTGAATCAAGGCTGTGTGCCAGCTGCTCAAC 446

DB 966 TTCTCCGCGGAGAAATTCGAGGTTCTGCAACGACCGACGTTCTCGCGCACCTGCTCGGC 907  
 QY 447 TGACGAGACCAATCAATGCACTGCTGTCAGGTATCGCGCGCTACTCTCTTGTCTGCG 506  
 DB 906 CCGGCTCGCGCGGATCAGCGGAGGTTCACGCGCG-----TTACG 865  
 QY 507 CAGCACCCATGCAAGCGCCATCTGCGCAAGTGAATGCGCGGTTCTTGGGCGATGCAAT 566  
 DB 864 CAGCACCCATGCAAGCGCCATCTGCGCAAGGCTCTGCGCGCGCTTCGCGCGATCGTATT 805  
 QY 567 GAGCTTGGGACCATATCAATATTGTTCAAGTTCAACATGCCCTCAGACAGGACTTACC 626  
 DB 804 GAGCTTGGCAGCTGCTCGAGGTTGTCGCGACTCAGGTGCTCTGCTTCAGGAGCGGCC 745  
 QY 627 -----CTGGCTGGCGCGGAAACCTCTGGAATTCATCGAGATATTTGTCGCTGAGCAG 680  
 DB 744 GCCCGGCTTGTTCACGCGCGGTGCGCGCGCGGTTTCAGGTACTTCGAGGTGAGCAA 685  
 QY 681 GCCCTGGCGCAAGTGTGAGAAAGCAATGACGCGCAAGACCATTTGTTGGCAGCTGACTGCAA 740  
 DB 684 GCCCTGGCGGAGCGCGCTGAACGCGATGCTGCGCGCGCGACGTCGTCGAGCGTACCGAG 625  
 QY 741 CAGGTTCTCACCG 753  
 DB 624 CAGATCCTCTTCG 612

Search completed: March 17, 2004, 09:47:12  
 Job time : 326.967 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:43:00 ; Search time 2400.75 Seconds  
(without alignments)  
10858.975 Million cell updates/sec

Title: US-09-105-117K-1\_COPY\_1421\_2293

Perfect score: 873

Sequence: 1 gfcgataagcagcgggttg.....ctgtgagctctggaccgtag 873

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*\*

2: em\_esthum:\*\*

3: em\_estin:\*\*

4: em\_estmu:\*\*

5: em\_estcov:\*\*

6: em\_estpl:\*\*

7: em\_estro:\*\*

8: em\_htc:\*\*

9: gb\_est1:\*\*

10: gb\_est2:\*\*

11: gb\_hrc:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estfun:\*\*

16: em\_estom:\*\*

17: em\_gss\_hum:\*\*

18: em\_gss\_inv:\*\*

19: em\_gss\_pln:\*\*

20: em\_gss\_vrt:\*\*

21: em\_gss\_fun:\*\*

22: em\_gss\_man:\*\*

23: em\_gss\_mus:\*\*

24: em\_gss\_pro:\*\*

25: em\_gss\_rhd:\*\*

26: em\_gss\_phg:\*\*

27: em\_gss\_vrl:\*\*

28: gb\_gss1:\*\*

29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	56.4	6.5	860	28 AF075981	AF075981 AF075981
c 2	54.2	6.2	897	28 BZ560550	BZ560550 pacs2-164
c 3	51.4	5.9	1620	28 BZ568946	BZ568946 pacs2-164
c 4	48.8	5.6	1101	29 CNS0175Y	AL108460 Drosophil

5	45	5.2	720	14	CD881613
6	44.4	5.1	664	29	CNS03K4J
7	43.8	5.0	3237	28	BH770958
8	43.2	4.9	1395	28	BZ574094
9	42.4	4.9	699	13	BX424825
10	42	4.8	457	13	BY246241
11	41.8	4.8	538	14	CD861975
12	41.6	4.8	558	12	BM712036
13	41.6	4.8	588	13	BU220119
14	41	4.7	617	12	BM729055
15	41	4.7	770	10	BE988495
16	41	4.7	861	10	BE740880
17	40	4.6	357	9	AI418751
18	40	4.6	431	9	AA452858
19	40	4.6	447	14	W91005
20	40	4.6	463	9	AA476500
21	40	4.6	504	9	AI159901
22	40	4.6	537	29	CG694695
23	40	4.6	540	9	AI401605
24	40	4.6	564	9	AI820014
25	40	4.6	600	12	BG806166
26	40	4.6	619	14	CF533294
27	40	4.6	656	28	BZ894552
28	40	4.6	3760	11	AK031268
29	40	4.6	5069	29	AY405421
30	40	4.6	5094	29	AY405420
31	39.8	4.6	1101	29	CNS00KK2
32	39.6	4.5	925	29	CNS0091P
33	39.4	4.5	546	14	CF879615
34	39.2	4.5	846	29	CNS010R3
35	39.2	4.5	1201	13	BX361080
36	39.2	4.5	1201	13	BX381961
37	39	4.5	543	14	CB640932
38	38.8	4.4	1201	13	BX381961
39	38.6	4.4	436	10	BE025218
40	38.6	4.4	513	13	BQ818660
41	38.6	4.4	828	29	CNS026XT
42	38.6	4.4	830	29	CNS01MQH
43	38.6	4.4	882	14	CF241805
44	38.6	4.4	1201	13	BX356664
45	38.6	4.4	1667	11	AY108197

# ALIGNMENTS

RESULT 1	AF075981/c	860 bp	DNA	linear	GSS 29-AUG-2000
LOCUS	AF075981	Salmonella typhimurium LT2	Lambda DASH II	Salmonella	
DEFINITION	AF075981	Salmonella typhimurium genomic clone 390-T3	genomic survey sequence		
ACCESSION	AF075981				
VERSION	AF075981.1	GI:3320851			
KEYWORDS	GSS				
SOURCE	Salmonella typhimurium				
ORGANISM	Salmonella typhimurium				
REFERENCE	1 (bases 1 to 860)				
AUTHORS	Wong, R.M.-Y., Wong, K.K., Benson, N.R. and McClelland, M.				
TITLE	Sample sequencing of a Salmonella typhimurium LT2 lambda library:				
JOURNAL	comparison to the Escherichia coli K12 genome				
MEDLINE	FEMS Microbiol. Lett. 173 (2), 411-423 (1999)				
PUBMED	99243757				
COMMENT	10227170				
	Contact: McClelland M				
	Molecular Biology				
	Sidney Kimmel Cancer Center				
	3099 Science Park Road, San Diego, CA 92121, USA				
	Email: mclelland@lifesci.sdsu.edu				
	Class: shotgun				
	Location/Qualifiers				
	1. .860				

FEATURES	
source	

```

/organism="Salmonella typhimurium"
/mol_type="genomic DNA"
/strain="L12"
/db_xref="taxon:602"
/clone="390-T3"
/notes="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer"

```

```

ORIGIN
Query Match      6.5%; Score 56.4; DB 28; Length 860;
Best Local Similarity 53.5%; Pred. No. 0.01;
Matches 199; Conservative 0; Mismatches 151; Indels 22; Gaps 3;

QY 505 CGCAGCACCCATCGAGCGCCATCTGCGCAAGTGACTGCGCGTCTCTGGCGCATGTCA 564
Db 400 CGCAGCATCCAGCGAGCGCCATTTGATACATTTTGGACCGTCTGGCGAGCAGTTCA 341
QY 565 TTGAGCTTTGGGACCATATCAATATTGTTTCACGTTCAACATCGCCCTCAGACAGGACTTA 624
Db 340 TTCAACCGCGAATCTTTTCCAGTTTGTGCGCGTAACTGTTCTGTGTTTAAAGGAACGG 281
QY 625 CCTCGCTG---GGCGGGAACCTCTGGAATTCATCGAATATTTGCTGCGTACGAGG 681
Db 280 CTTCCGCTCGCGCGCGGGAATCTTCCGGAATACCAATTCAAATTAACGTCAGTGAGC-TG 222
QY 682 CCTCGCAAGTGTGAGAAAGCAATGACGCCAAGACCAATTGTTGGCAGCTGACTGCAAC 741
Db 221 CCGCGCCACCGCGAGAGGCAATACTACGACGCTTTTCTCTACACAGCCGACG 162
QY 742 AAGTCTTCACCGTCATGCGCGTTCCTCCACCAACGATTAATGATGGAATAGTTGCG 801
Db 161 AG-----CCGCTCTCTACCAACGCTCAAAAAGCGAATATTATAGC 120
QY 802 TGATGAATCAGACCGCGGAGCGCTCTCCGCGCATCACTCAGCGCTCGCTGAGC 861
Db 119 TGATGAATCAGGCAAGCGTGGCGGAGATCTCCAGAAATATGATAGCTTCTTGCCGAGA 60
QY 862 TCTGGACCGTAG 873
Db 59 TCGGACGATAG 48

```

```

RESULT 2
BZ560550/c
LOCUS      897 bp      DNA      linear      GSS 17-DEC-2002
DEFINITION pacs2-164_2569.xl pacs2-164 Pseudomonas aeruginosa genomic clone
ACCESSION  BZ560550
VERSION    BZ560550.1
KEYWORDS   GSS.
SOURCE     Pseudomonas aeruginosa
ORGANISM   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE  1 (bases 1 to 897)
AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE     Whole-Genome-Sequence variation among multiple isolates of
JOURNAL   Pseudomonas aeruginosa library
COMMENT   J. Bacteriol. (2002) in press
            Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES   Location/Qualifiers
            source          1..897
                        /organism="Pseudomonas aeruginosa"
                        /mol_type="genomic DNA"

```

```

/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_2569"
/clone_lib="pacs2-164"
/notes="clinical isolate 2-164 Whole genomic shotgun
library."

```

```

ORIGIN
Query Match      6.2%; Score 54.2; DB 28; Length 897;
Best Local Similarity 52.8%; Pred. No. 0.035;
Matches 143; Conservative 0; Mismatches 120; Indels 8; Gaps 1;

QY 35 TGATGGCAATCGTGTGACTGCTGTTGAACCGGAATCGGTATTTGACGCGTTGTGTTTA 94
Db 552 TGGCGGCACTCGGCGTTCACCTGGCTTAAACCCCAACGCTCTATTTCGATACCGTACTTCTTAC 493
QY 95 TCGGGGCGGTGCGGCGCAATACGCGCACACCGGCGGTGGAATTTTCGCGCTGCGCGGT 154
Db 492 GGTTCGTCGCGCGCAGCAAGNCCGCG-----CCGGCGCATATGCCCTCGGCGCG 441
QY 155 TCGCGCGAAGCTGATCTGTTCCCGCTGCGTGGTGGTTTCGCGCAGCAGCATTTGTACGCC 214
Db 440 CCAAGCGCTCGCTGATGTGTTCTTCGCGCTCGCCCTCGGCGCGCATGGCTGGCCCCCT 381
QY 215 CGCTGTCAGCCCCCAAGGTGTGCGCTGATCAACGTCGTCGCGAGTTGTGATGACCG 274
Db 380 GGTGCGCGCGCGCGCCACCTGGCGCTGTCGACCTGATGTGCGCGCCATGATGCTGG 321
QY 275 CATTGGCCATCAAACTGATGTTGATGGTTA 305
Db 320 GCATGGCGCGCAACTGCTGTTCCGGGATA 290

```

```

RESULT 3
BZ568946/c
LOCUS      1620 bp      DNA      linear      GSS 17-DEC-2002
DEFINITION pacs2-164_8165.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
ACCESSION  BZ568946
VERSION    BZ568946.1
KEYWORDS   GSS.
SOURCE     Pseudomonas aeruginosa
ORGANISM   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE  1 (bases 1 to 1620)
AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE     Whole-Genome-Sequence variation among multiple isolates of
JOURNAL   Pseudomonas aeruginosa library
COMMENT   J. Bacteriol. (2002) in press
            Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES   Location/Qualifiers
            source          1..1620
                        /organism="Pseudomonas aeruginosa"
                        /mol_type="genomic DNA"
                        /strain="2-164"
                        /db_xref="taxon:287"
                        /clone="pacs2-164_8165"
                        /clone_lib="pacs2-164"
                        /notes="clinical isolate 2-164 Whole genomic shotgun
                        library."
ORIGIN
Query Match      5.9%; Score 51.4; DB 28; Length 1620;
Best Local Similarity 54.3%; Pred. No. 0.22;

```

	Query Match	5.6%	Score 48.8	DB 29	Length 1101
	Best Local Similarity	14.5%	Pred. No. 0.78		
	Matches	47	Conservative 156	Mismatches 121	Indels 0
					Caps 0
QY	72	GTATTGACGCGCTTTGTTTATCTCGCGCGCTCGCGCGCAATACGGCGACACCGGACG	131		
DB	1079	KTTTATTBSSBSTSTSTBTBBSBSBBBSBSBSBSBSBSBSSTSSSSSSSTSSSS	1020		
QY	132	GTGGAATTTCCGCGTGGCGGCTGTCGCGGACGCTGATCGCTTCCGCTGTGGGTTT	191		
DB	1019	STSBSTBTSTSSSSTTTTTTTTTTTSTBTBBSBSBSBSBSBSSTSSSBSTSKBTSBS	960		
QY	192	CGCGCGACGACGATTCACGCGCCGCTGTCCAGCCCAAGGTGTGGCGCTGGATCAAGCT	251		
DB	959	BGTSSAGBSBSBSBSBSSTBTASTBBSBTSSTSSGSGSSSSBSBSTBSBSBTBT	900		

Qy	1	GTCGATAAGCAGCGGGTTTGGGTAAGAGCCCATGTTTGATATGCAATCGTGCTGACCTGGTTG	60
Db	252	GTGCGAGAGAGACTGTGTGGCGCTCAAGCCCAAGAGCATCGACTTTCGCGAGGCGCGCGGT	311
Qy	61	AACCCGAATCGGTATTTGAGCGCGTTTGTTGTTATTCGCGCGCGTGGGGCGCAATACGC	120
Db	312	CTGCCGCTCCGCAATCGAGACGCCCATGAGGCTTCGAGAGGCGGGCTTCTCCGCGCGG	371
Qy	121	GACACCGGAGCGTGGATTTTCGCCGCTGGCGGTTTCGGCGAAGCCTGATCTGTGTTCCCG	180
Db	372	AAGTCATCTCTGCTCTCGCGCGGCGCGCGAGTTCGGGACCCCTGTCATCCAGCTGGG	431
Qy	181	CTGTGGGTGTTTCGGCGCAGCAGCATTTGTACAGCCCGCTGTTCAGGCCCAAGGTGTGGCG	240
Db	432	AGCGAAGTTTACGGCGCTTCACGCTGGGGGCCACGCGCAGACCCCGAAGCTGGAGCTC	491
Qy	241	TGGATCAACGCTCGTGGCGAGTTTGAT	269
Db	492	CTGAAAGCCTGGGAGCCGACGCTGCCCAT	520

RESULT 6  
CNS03K4J  
LOCUS  
DEFINITION  
CNS03K4J 664 bp DNA linear GSS 01-SEP-2000  
Tetraodon nigroviridis genome survey sequence T7 end of clone  
032806 of library G from Tetraodon nigroviridis, genomic survey  
sequence.  
AL247708  
GSS; genome survey sequence.  
Tetraodon nigroviridis  
Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontoidea; Tetraodontidae; Tetraodon.  
1  
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,  
Bernot,A., Fzames,C., Wincker,P., Brottier,P., Quecuer,F.,  
Saurin,W. and Weissenbach,J.  
Estimate of human gene number provided by genome-wide analysis  
using Tetraodon nigroviridis DNA sequence  
Nat. Genet. 25 (2), 235-238 (2000)  
20296633  
10835645  
2  
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,  
Bernot,A., Fzames,C., Wincker,P., Brottier,P., Quecuer,F.,  
Saurin,W. and Weissenbach,J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis  
Genome Res. 10 (7), 939-949 (2000)  
20359837  
10899143  
3 (bases 1 to 664)  
Genoscope.  
Direct Submission  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetraodon nigroviridis  
genome. For more information, please take a look at  
http://www.genoscope.cns.fr/Tetraodon.  
FEATURES  
source  
1..664  
Location/Qualifiers  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="032806"  
/clone\_lib="G"  
/note="Genoscope sequence ID : C0B6032DA031P1-end : T7"

Query Match 5.1%; Score 44.4; DB 29; Length 664;  
Best Local Similarity 48.2%; Pred. No. 7;  
Matches 120; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 181 CTGGTGGGTTTCGGCGCAGCAGCATTTGTCAGCCCGCTGTCCAGCCCAAGGTTGGCGC 240  
DB 278 CAGGTCCGCTCTCGCCCGGAAAATGCGGGGCCCGCGGAGCGCCGCGGCGGCGC 337  
QY 241 TGATCAACGTCGTGGCAGTTGTGATGACCGCATTTGGCCATCAACTGATTTGATG 300  
DB 338 CGCGGNTCTTTTCTGCGGCGGTGGCGCGGACAGGCCCGCGCCGCTGNTGGTGGT 397  
QY 301 GGTAGTTTTCGGGGTTTGGAAATCGGTGGCCTTTCGCCCAAAATGTTGATGCCGGGTG 360  
DB 398 CGGACGATCCGGTGGGTCGGGASCCGTGGCTCTCGCACCGATGGAGGCCCTGGGGCG 457  
QY 361 TGGGAAATCTCATGATCGCCTCCAACTCGGCGTCCAGAAATCTCAAGTTGTTGAGTGAA 420

DB 458 GAGGCTCCCCACGCTGGGAGCTCCCGCGCGTGGAGGCTCCCGCGTGGTGGGCTCTT 517  
QY 421 TCNAGGCTG 429  
DB 518 CCCCGCGCG 526

RESULT 7  
BH770958 3237 bp DNA linear GSS 01-MAY-2002  
LOCUS  
DEFINITION  
LIM2tag686 MG1363 Random Sequence Tag Library Lactococcus lactis  
subsp. cremoris genomic, genomic survey sequence.  
BH770958  
GSS.  
BH770958.1 GI:20373915  
Lactococcus lactis subsp. cremoris  
Lactococcus lactis subsp. cremoris  
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
Lactococcus.  
1 (bases 1 to 3237)  
Bolotin,A., Ehrlich,S.D. and Sorokin,A.  
Studies of genomes of dairy bacteria Lactococcus lactis  
Sci. Aliments (2002) In press  
Contact: Sorokin A  
Genetique Microbienne  
INRA  
CRU INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France  
Tel: 33 1 34 65 25 16  
Fax: 33 1 34 65 25 21  
Email: sorokine@jouy.inra.fr  
best homologue in strain IL1403 is nadR (98%)  
Class: shotgun  
High quality sequence start: 30  
High quality sequence stop: 3207.  
Location/Qualifiers  
1..3237  
/organism="Lactococcus lactis subsp. cremoris"  
/mol\_type="genomic DNA"  
/strain="MG1363"  
/sub\_species="cremoris"  
/db\_xref="taxon:1359"  
/clone\_lib="MG1363 Random Sequence Tag Library"  
/note="Vector: PSGM2; Site 1: SmaI; Library of  
chromosomal fragments of L.lactis strain MG1363 was  
prepared by partial AluI digestion or by sonication."

FEATURES  
source  
1..3237  
Location/Qualifiers  
/organism="Lactococcus lactis subsp. cremoris"  
/mol\_type="genomic DNA"  
/strain="MG1363"  
/sub\_species="cremoris"  
/db\_xref="taxon:1359"  
/clone\_lib="MG1363 Random Sequence Tag Library"  
/note="Vector: PSGM2; Site 1: SmaI; Library of  
chromosomal fragments of L.lactis strain MG1363 was  
prepared by partial AluI digestion or by sonication."

ORIGIN  
Query Match 5.0%; Score 43.8; DB 28; Length 3237;  
Best Local Similarity 49.9%; Pred. No. 22;  
Matches 170; Conservative 0; Mismatches 162; Indels 9; Gaps 2;

QY 385 AACTCGGCTCAGAAAATCCAGTTGTTGAGTGAATCAAGCTGTGTCAGCTGTCTCA 444  
DB 1956 AATCTTCGTAGTAATCCAGAGTTTAAAGCTGATTAACAGTTGAACCTTTGCCAGCTCTT---TCC 2015  
QY 445 ACTGACGAGCAGCAACCAATCAATGCTGCTACGGTATCCGCGCCGCTACTCTCTGCTCG 504  
DB 2016 GGACGACTTGGCCCAATCAAGCTGATTAACAGTTGAACCTTTGCCAGCTCTT---TCC 2072  
QY 505 CGGACGCCCATTCGACAGCGGCTGTCGCGAAGTCACTGCCCGTCTCGGCGATGTCA 564  
DB 2073 CGTAAACCAAGCTTATGCACTTTGAGCTAGATTTGTCACAGCTTTGAGCTTAATCA 2132  
QY 565 TTGAGCTCTGGGACCATATCAATATTGTTTCAGCTTCAATGCTCCCTCAGACAGGACTTA 624  
DB 2133 TTAGGGCTTGAACCTGTTTCCAATCTTTCTGTGGTAGACTGTCAATCAAGTCGA 2192  
QY 625 CCCTGGCTGGCGCG---GAAACCCCTCTGGAATTCATCCAGATATTGTCGAGC 678  
DB 2193 TAATGAGGATCAGCATTCGCGAATTTTCAGGAATTCATGCAATAATTTCTCTGTTAAG 2252  
QY 679 AGGCCCTGGCGAAGTGGTGAAGCAATGAGCCGACGACC 719



Db 2253 AGTCCTTGATCAAAAGGCTTAAGGCAATGTGTTCTATCCC 2293

## RESULT 8

BZ574094/c

## LOCUS

DEFINITION BZ574094 1395 bp DNA linear GSS 17-DEC-2002  
msh2\_3508.y3 msh Pseudomonas aeruginosa genomic clone msh2\_3508,  
genomic survey sequence.

## ACCESSION

BZ574094

## VERSION

BZ574094.1 GI:27209155

## KEYWORDS

GSS.

## SOURCE

Pseudomonas aeruginosa

## ORGANISM

Pseudomonas aeruginosa

## REFERENCE

1 (bases 1 to 1395)

## AUTHORS

Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,  
Burns,J.L., Kaul,R. and Olsen,M.V.

## TITLE

Whole-Genome-Sequence variation among multiple isolates of

## JOURNAL

Pseudomonas aeruginosa library

## COMMENT

J. Bacteriol. (2002) In press

## CONTACT

Contact: Chris K. Raymond

## Genome Center

University of Washington

## Box

352145, Seattle, WA 98105-2145, USA

## Tel:

2062216954

## Fax:

2066857244

## Email:

craymond@u.washington.edu

## Class:

shotgun.

## Location/Qualifiers

1..1395

## /organism="Pseudomonas aeruginosa"

/mol\_type="genomic DNA"

## /strain="GSH"

/db\_xref="taxon:287"

## /clone="msh2\_3508"

/note="Environmental isolate. Whole genomic shotgun  
library."

## ORIGIN

Query Match 4.9%; Score 43.2; DB 28; Length 1395;

## Best Local Similarity

55.3%; Pred. No. 20;

## Matches

84; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

## QY

488 CGTACTCTCTGCTCGCGGAGCACCACGAGCCATCGAGCCATCGGCAAGTCACTGCCCGC 547

## Db

527 CGTCTCGGCTTGTCACCGAGCACCAGCGGAGCACCATCGGGGGGATGGGTAACCGA 468

## QY

548 GTTCCTGGGCGATGTCATGAGCTTGCGGACCATATCAATATTGTTCACTTCAACATGC 607

## Db

467 GTTCTTGTAACAGGTAGCGGAGCGCGGACAGCAGCATCAGGACACGTCGAGCATGG 408

## QY

608 CTTAGCAGGAGGACTTACCTCGGCTGGCGCGG 639

## Db

407 CGTTGTGACCGAGTAGGCTCCGATGGCGCAG 376

## RESULT 9

BX424825/c

## LOCUS

DEFINITION BX424825 Homo sapiens PLACENTA Homo sapiens CDNA clone

## ACCESSION

BX424825

## VERSION

BX424825.1 GI:30647817

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 699)

## AUTHORS

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

## TITLE

Full-length cDNA libraries and normalization

## JOURNAL

## COMMENT

Unpublished (2001)

## Contact:

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Paraday Avenue Genoscope sequence ID: XCL0BA001Z02P01.

## FEATURES

Location/Qualifiers

1..699

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="XCL0BA001Z02"

/tissue\_type="PLACENTA"

/clone\_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT.6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

## ORIGIN

Query Match 4.9%; Score 42.4; DB 13; Length 699;

Best Local Similarity 16.4%; Pred. No. 22;

Matches 44; Conservative 113; Mismatches 112; Indels 0; Gaps 0;

## QY

95 TCGGGCGGCTCGGGCGCAATACGGGACACCGAGGTGGATTTCGCGCTGGCGCGT 154

## Db

670 BCGGG 611

## QY

155 TCGGGCGCAAGCTGATCTGTTCCCGCTGGTGGTTCGGCGCAGCAGCATTTGTCAGCC 214

## Db

610 CCKKKKGGGKKKCCCCCCCCCKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 551

## QY

215 CGCTGTCCAGCCCGCAGGTGGCGCTGGATCAAGCTGCTGGGAGCATTTGATGACCG 274

## Db

550 KKK 491

## QY

275 CATTTGGCCATCAACTGATGTTGATGGTGTGATGTTTCGCGGGTTTGGAAATCGTGGCGCT 334

## Db

490 KKK 431

## QY

335 TCGGCCAAATGTTGATGCGCGGCTGGTGG 363

## Db

430 CCGCCCCCKKK 402

## RESULT 10

BY246241/c

## LOCUS

DEFINITION BY246241 457 bp mRNA linear EST 10-DEC-2002

CDNA clone K230304B15 5', mRNA sequence.

## ACCESSION

BY246241

## VERSION

BY246241.1 GI:26427753

## KEYWORDS

EST.

## SOURCE

Mus musculus (house mouse)

## ORGANISM

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 457)

## REFERENCE

Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,

Nikaido,I., Osato,N., Saito,R., Yamazaki,H., Yamazaki,I.,

Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,

Schönbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,

Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,

Batalov,S., Bessel,K.W., Blake,J.A., Bradt,D., Brusic,V.,

Nikolich,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,

Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,

Gariboldi,M., Gissi,C., Godzik,J., Gough,J., Grimmond,S.,

Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,

Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,

Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., Mckenzie, D., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Read, J.C., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynnaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, K., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354583

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp

URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, T., Itoh, M., Kawai, J., Konno, H., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ono, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Michela Pagioli and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hiroawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

TITLE

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

200 CTTGTCCTCGGACAGAAAGCTCAGTCTGGTGGAGACAGGCTGGTATTCGGAC 141  
440 GGTCACTGACGAGACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 499  
140 CTTGTCCTCGGACAGAAAGCTCAGTCTGGTGGAGACAGGCTGGTATTCGGAC 81  
500 GCTGCGCAGCAGCCATCAAGCCCATCTGGGCAAGTGAATGCTGCGCGCTTCTCTGGGCGA 559  
80 CATGCGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 21  
560 TGTGAT 565  
20 CGGCT 15

CD861975 538 bp mRNA linear EST 11-JUL-2003  
AZ01.101N14F010125 AZ01 Triticum aestivum cDNA clone AZ01101N14,  
mRNA sequence.

CD861975  
CD861975.1 GI:32545791  
Triticum aestivum (bread wheat)  
EST.  
Triticum aestivum  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
1 (bases 1 to 538)  
Genoplante.  
Genoplante, a major partnership french program in plant genomics  
Unpublished (2003)  
Contact: Genoplante  
Genoplante  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french  
plant genomics programme 'Genoplante' (http://www.genoplante.com  
and http://genoplante-info.infobiogen.fr).

FEATURES  
source  
1..538  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="recital"  
/db\_xref="taxon:4565"  
/clone="AZ01101N14"  
/tissue\_type="leaf"  
/clone\_lib="AZ01"

ORIGIN  
Query Match 4.8%; Score 41.8; DB 14; Length 538;  
Best Local Similarity 47.2%; Pred. No. 27;  
Matches 127; Conservative 0; Mismatches 142; Indels 0; Gaps 0;  
QY 1 GTGCAATAGCAGCGGTTTGGTAAAGCCCATGTTGATGCGCAATCGTGTGACCTGGTTG 60  
109 GTGCGAGAGAGCTGTGGTGGCTCAAGCCCAAGAGCATCGACTTCGGCGCAGCGCGCGT 168  
61 AACCCGATGCTATTGAGCGGTTGTGTTTATCGCGCGGTGGCGGCAATACGCG 120  
169 CTGCGCGCTCGCCATCGAGAGACCGCCCATGAGGCGCTCGAGAGAGAGGCTTCTCCGCGCG 228  
121 GACACCGGAGCGGTGGATTTTCGCGCGCTGCGCGCGCTGCGCGCAAGCTGATCTGTTCCCG 180  
229 AAGTCCATCTTGTCTCGCGCGCGCGCGCGGAGTGGGAGCCCTTGTTCATCCAGCTGGCA 288  
181 CTGCTGGGTTTGGCGCGAGCAGATTGTGACGCGCGCTGTCAGCCCAAGAGTGGCGCGC 240  
289 AAGCAAGTTTACGGCGCATCGAAAGTGGCGCGCGCACAGCAGCAGCCCAAGAGCTGGAGTC 348  
241 TGGATCAGCTGCTGTGGGAGCTGTGAT 269

ORIGIN

Query Match 4.8%; Score 42; DB 13; Length 457;  
Best Local Similarity 51.6%; Pred. No. 22;  
Matches 96; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
QY 380 CTTCCAACTCGGCGTCAGAAACTCAAGTTGTGTGATGATCAAGGCTGTGTCAGCT 439







GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2004, 18:03:00 ; Search time 60 Seconds  
(without alignments)  
1111.354 Million cell updates/sec

Title: US-09-105-117K-2

Perfect score: 1191  
Sequence: 1 MWMEIFITGLLGASLLS.....INVVAVVMTALAIXIMMG 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 segs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001s.\*

5: Geneseq2002s.\*

6: Geneseq2003as.\*

7: Geneseq2003bs.\*

8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1191	100.0	236	2	AAW37715 C. glutam
2	1191	100.0	236	4	AAW37715 C. glutam
3	1191	100.0	236	6	ABR58213 LysE prot
4	1178	98.9	233	4	AAW79660 Coryneb
5	1178	98.9	233	4	AAU71888 C. glutam
6	1178	98.9	233	7	ADB66197 protein e
7	864.5	72.6	231	4	AAW64047 Coryneb
8	584	49.0	124	6	ABR58214 LysE24 pr
9	350	29.4	211	3	AAW01789 Escherich
10	298	25.0	208	6	ABM68839 Phototrab
11	298	25.0	211	6	ADA34450 Acinetoba
12	170.5	14.3	210	2	AAW20426 H. pylori
13	170.5	14.3	215	2	AAW20968 H. pylori
14	164	13.8	205	6	ABM72206 Staphyloc
15	132.5	11.1	205	6	ABM67507 Phototrab
16	132.5	11.1	226	4	AAW89911 C. glutam
17	121.5	10.2	235	6	ADA36962 Acinetoba
18	119.5	10.0	173	5	ABP39181 Staphyloc
19	119.5	10.0	222	6	ADA35184 Acinetoba
20	119	10.0	620	6	ABU49694 Protein e
21	114.5	9.6	229	3	ABU1637 A. vitis
22	112	9.4	203	6	ABM70439 Phototrab
23	108.5	9.1	222	6	ABM68944 Phototrab
24	108.5	9.1	238	6	ADA35835 Acinetoba
25	105	8.8	206	3	AAW99598 E. coli L

## ALIGNMENTS

### RESULT 1

AAW37715  
ID AAW37715 standard; protein; 236 AA.

XX AC AAW37715;

DT 12-MAR-1998 (first entry)

XX C. glutamicum Lys E protein (lysine export protein).

XX LysG; LysE; ORF3; lysine transport; regulatory protein; export protein;  
XX Microbial production; amino acid; animal feed additive.

XX Corynebacterium glutamicum.

XX DE19548222-A1.

XX 26-JUN-1997.

XX 22-DEC-1995; 95DE-01048222.

XX 22-DEC-1995; 95DE-01048222.

XX (KERJ ) FORSCHUNGSZENTRUM JUELICH GMBH.

XX Vrijic M, Eggeling L, Sahm H;

XX WPI; 1997-333867/31.

XX N-PSDB; AAU96816.

PT Increasing microbial production of amino acids, especially lysine - by  
PT improving export carrier activity or corresponding gene expression, also  
PT new export and regulatory genes from Corynebacterium.

XX Claim 42; Page 10; 16pp; German.

XX This sequence is the LysE protein product, a lysine export protein. LysG  
XX and LysE encode a lysine transport regulatory protein and an export  
XX protein, respectively. Microbial production of amino acids (A) is  
XX improved by increasing the export-carrier activity and/or the export gene  
XX expression in a microorganism that produces (A). The method is  
XX specifically used to increase production of lysine, used as an animal  
XX feed additive. Other (A) are variously useful as pharmaceuticals,  
XX condiments and intermediates for fine chemicals. This method increases  
XX the amount of (A) secreted into the culture medium. Export of (A) has  
XX been found to depend on a single gene

XX Sequence 236 AA;

ABU16949 Protein e  
ADA33616 Acinetoba  
ABM68999 Phototrab  
AAB01787 Escherich  
ADA33824 Acinetoba  
ABM67782 Phototrab  
ABU49858 Protein e  
AAU58229 Propionib  
ABM54748 Propionib  
ADA36250 Acinetoba  
AAU34590 E. coli c  
ABU28648 Protein e  
ABU19902 Protein e  
AAB78987 C. glutam  
AAG93244 C. glutam  
AAB01786 Escherich  
ABB93711 Herbicida  
AAB76722 Coryneb  
AAG92602 C. glutam  
AAE22703 Coryneb

26 103.5 8.7 181 6 ABU16949  
27 101.5 8.5 220 6 ADA33616  
28 101 8.5 201 6 ABM68999  
29 101 8.5 212 3 AAB01787  
30 99 8.3 208 6 ADA33824  
31 96 8.1 213 6 ABM67782  
32 95 8.0 295 6 ABU49858  
33 95 8.0 390 4 AAU58229  
34 95 8.0 390 6 ABM54748  
35 94.5 7.9 240 6 ADA36250  
36 93.5 7.9 452 4 AAU34590  
37 93.5 7.9 452 6 ABU28648  
38 93 7.8 211 6 ABU19902  
39 93 7.8 459 4 AAB78987  
40 93 7.8 459 4 AAG93244  
41 92 7.7 223 3 AAB01786  
42 92 7.7 694 5 ABB93711  
43 91.5 7.7 446 4 AAB76722  
44 91.5 7.7 446 4 AAG92602  
45 91.5 7.7 446 5 AA E22703

Query Match 100.0%; Score 1191; DB 2; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-128; Indels 0; Gaps 0;  
 Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWMEIFITGLLGGASLLSIGPQNVVVKQIKREGLIAVLLVCLISDVFLFIAGTGLV 60  
 |||||  
 DB 1 MWMEIFITGLLGGASLLSIGPQNVVVKQIKREGLIAVLLVCLISDVFLFIAGTGLV 60  
 |||||

QY 61 DLLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTKVAPQIIEETPTVPDDTPLGGS 120  
 |||||  
 DB 61 DLLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTKVAPQIIEETPTVPDDTPLGGS 120  
 |||||

QY 121 AVATDRNRVRVEVDKQVWVKPMLMAIVLTWLPNAYLDFAFVIGGGAQYGTGRW 180  
 |||||  
 DB 121 AVATDRNRVRVEVDKQVWVKPMLMAIVLTWLPNAYLDFAFVIGGGAQYGTGRW 180  
 |||||

QY 181 IFAGAFASLIWPLVGFGAALSRPLSSPKVWRINVVAVVMTALAIKMLMG 236  
 |||||  
 DB 181 IFAGAFASLIWPLVGFGAALSRPLSSPKVWRINVVAVVMTALAIKMLMG 236  
 |||||

## RESULT 2

AAG93201  
 ID AAG93201 standard; protein; 236 AA.  
 AC AAG93201;  
 DT 26-SEP-2001 (first entry)  
 DE C glutamicum protein fragment SEQ ID NO: 6955.  
 DE Corynebacterium; amino acid synthesis; vitamin; saccharide;  
 KW organic acid synthesis.  
 KW Corynebacterium glutamicum.  
 OS EP1108790-A2.  
 PN 20-JUN-2001.  
 PD 18-DEC-2000; 2000EP-00127688.  
 PF 16-DEC-1999; 99JP-00377484.  
 PR 07-APR-2000; 2000JP-00159162.  
 PR 03-AUG-2000; 2000JP-00280988.  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 XX WPI: 2001-376931/40.  
 DR N-PSDB; AAH68420.  
 XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analyzing  
 PT expression profile or pattern of a gene and identifying homologous gene.  
 XX Claim 29; SEQ ID NO 6955; 246pp + Sequence Listing; English.  
 XX The present invention provides a number of nucleotide and protein  
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of corynebacterium bacterium, measuring expression amount and analysing  
 CC the expression profile or expression pattern of a gene derived from  
 CC Corynebacterium bacterium, and identifying a homologue of a gene derived from  
 CC Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino  
 CC acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a protein described in the  
 CC exemplification of the invention. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from the European Patent Office

XX SQ Sequence 236 AA;  
 Query Match 100.0%; Score 1191; DB 4; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-128; Indels 0; Gaps 0;  
 Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWMEIFITGLLGGASLLSIGPQNVVVKQIKREGLIAVLLVCLISDVFLFIAGTGLV 60  
 |||||  
 DB 1 MWMEIFITGLLGGASLLSIGPQNVVVKQIKREGLIAVLLVCLISDVFLFIAGTGLV 60  
 |||||

QY 61 DLLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTKVAPQIIEETPTVPDDTPLGGS 120  
 |||||  
 DB 61 DLLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTKVAPQIIEETPTVPDDTPLGGS 120  
 |||||

QY 121 AVATDRNRVRVEVDKQVWVKPMLMAIVLTWLPNAYLDFAFVIGGGAQYGTGRW 180  
 |||||  
 DB 121 AVATDRNRVRVEVDKQVWVKPMLMAIVLTWLPNAYLDFAFVIGGGAQYGTGRW 180  
 |||||

QY 181 IFAGAFASLIWPLVGFGAALSRPLSSPKVWRINVVAVVMTALAIKMLMG 236  
 |||||  
 DB 181 IFAGAFASLIWPLVGFGAALSRPLSSPKVWRINVVAVVMTALAIKMLMG 236  
 |||||

## RESULT 3

ABRS8213  
 ID ABRS8213 standard; protein; 236 AA.  
 AC ABRS8213;  
 DT 23-OCT-2003 (revised)  
 DT 11-AUG-2003 (first entry)  
 DE LysE protein.  
 DE L-lysine; L-arginine; LysE.  
 KW Corynebacterium glutamicum.  
 OS EP1266966-A2.  
 PN 18-DEC-2002.  
 PD 05-JUN-2002; 2002EP-00012539.  
 PF 12-JUN-2001; 2001JP-00177075.  
 PR (AJIN ) AJINOMOTO CO INC.  
 XX Gunji Y, Yasueda H;  
 PI WPI: 2003-241171/24.  
 DR N-PSDB; ACC80941.  
 XX Novel DNA encoding variant of LysE protein from a coryneform bacterium,  
 PT when introduced into methanol assimilating bacterium, facilitates  
 PT excretion of L-lysine and/or L-arginine to outside of a cell.  
 XX Disclosure; Page 18-19; 23pp; English.  
 XX The present invention relates to DNA encoding variants of protein with  
 CC loop region and six hydrophobic helices which facilitates excretion of L-  
 CC lysine and/or L-arginine to outside of cell of a methanol assimilating  
 CC bacterium when introduced into the bacterium. The method is used for  
 CC encoding a protein which facilitates excretion of L-lysine, L-arginine or  
 CC both of these L-amino acids to outside of a cell of a methanol  
 CC assimilating bacterium when DNA of the method is introduced into the  
 CC bacterium. The present sequence represents a lysE protein from  
 CC Brevibacterium lactofermentum. (Updated on 23-OCT-2003 to standardise OS  
 CC field)  
 XX SQ Sequence 236 AA;



Query Match 100.0%; Score 1191; DB 6; Length 236;  
Best Local Similarity 100.0%; Pred. No. 4.6e-128;  
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYWEIPIITGLLGASLLSIPQNVIVIKQIKREGIAVLVCLISDVFLFIAGTGLV 60  
DB 1 MYWEIPIITGLLGASLLSIPQNVIVIKQIKREGIAVLVCLISDVFLFIAGTGLV 60

QY 61 DILSNAPIVLDIMRWGGIAYLLWFAVMAAKDANTNKVEAPQIIEETPTVDDTPLGGS 120  
DB 61 DILSNAPIVLDIMRWGGIAYLLWFAVMAAKDANTNKVEAPQIIEETPTVDDTPLGGS 120

QY 121 AVATDTRNRVREVSVDKQVWVKPMLMAVLITMLNPNAYLDAFVFIGVGAQYGDGTGRW 180  
DB 121 AVATDTRNRVREVSVDKQVWVKPMLMAVLITMLNPNAYLDAFVFIGVGAQYGDGTGRW 180

QY 181 IPAAGAFASLIWFLVGFGAALSRPLSPKVRWVNVVAVVMTALAKMLMG 236  
DB 181 IPAAGAFASLIWFLVGFGAALSRPLSPKVRWVNVVAVVMTALAKMLMG 236

RESULT 4  
AAB79660  
ID AAB79660 standard; protein; 233 AA.  
AC AAB79660;  
XX  
XX  
XX 30-APR-2001 (first entry)  
XX  
XX Corynebacterium glutamicum MP protein sequence SEQ ID NO:54.  
XX  
XX Corynebacterium glutamicum; metabolic pathway protein; MP protein;  
KW fine chemical production; microorganism; organic acid; nucleoside;  
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;  
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.  
XX  
XX Corynebacterium glutamicum.  
XX  
XX WC200100843-A2.  
XX  
XX 04-JAN-2001.  
XX  
XX 23-JUN-2000; 2000WO-IB000923.  
XX  
XX 25-JUN-1999; 99US-0141031P.  
PR 01-JUL-1999; 99DE-01030476.  
PR 02-JUL-1999; 99US-0142101P.  
PR 08-JUL-1999; 99DE-01031415.  
PR 08-JUL-1999; 99DE-01031418.  
PR 08-JUL-1999; 99DE-01031419.  
PR 08-JUL-1999; 99DE-01031420.  
PR 08-JUL-1999; 99DE-01031424.  
PR 08-JUL-1999; 99DE-01031428.  
PR 08-JUL-1999; 99DE-01031434.  
PR 08-JUL-1999; 99DE-01031435.  
PR 08-JUL-1999; 99DE-01031443.  
PR 08-JUL-1999; 99DE-01031453.  
PR 08-JUL-1999; 99DE-01031457.  
PR 08-JUL-1999; 99DE-01031465.  
PR 08-JUL-1999; 99DE-01031478.  
PR 08-JUL-1999; 99DE-01031510.  
PR 08-JUL-1999; 99DE-01031541.  
PR 08-JUL-1999; 99DE-01031541.  
PR 08-JUL-1999; 99DE-01031573.  
PR 08-JUL-1999; 99DE-01031592.  
PR 08-JUL-1999; 99DE-01031634.  
PR 08-JUL-1999; 99DE-01031632.  
PR 08-JUL-1999; 99DE-01031634.  
PR 08-JUL-1999; 99DE-01031636.  
PR 09-JUL-1999; 99DE-01032125.  
PR 09-JUL-1999; 99DE-01032126.  
PR 09-JUL-1999; 99DE-01032130.  
PR 09-JUL-1999; 99DE-01032186.  
PR 09-JUL-1999; 99DE-01032206.

PR 09-JUL-1999; 99DE-01032227.  
PR 09-JUL-1999; 99DE-01032228.  
PR 09-JUL-1999; 99DE-01032229.  
PR 09-JUL-1999; 99DE-01032230.  
PR 14-JUL-1999; 99DE-01032922.  
PR 14-JUL-1999; 99DE-01032926.  
PR 14-JUL-1999; 99DE-01032928.  
PR 14-JUL-1999; 99DE-01033004.  
PR 14-JUL-1999; 99DE-01033005.  
PR 14-JUL-1999; 99DE-01033006.  
PR 12-AUG-1999; 99US-0148613P.  
PR 27-AUG-1999; 99DE-01040764.  
PR 27-AUG-1999; 99DE-01040765.  
PR 27-AUG-1999; 99DE-01040766.  
PR 27-AUG-1999; 99DE-01040832.  
PR 31-AUG-1999; 99DE-01041378.  
PR 31-AUG-1999; 99DE-01041379.  
PR 31-AUG-1999; 99DE-01041380.  
PR 31-AUG-1999; 99DE-01041394.  
PR 31-AUG-1999; 99DE-01041396.  
PR 03-SEP-1999; 99DE-01042076.  
PR 03-SEP-1999; 99DE-01042077.  
PR 03-SEP-1999; 99DE-01042079.  
PR 03-SEP-1999; 99DE-01042086.  
PR 03-SEP-1999; 99DE-01042087.  
PR 03-SEP-1999; 99DE-01042088.  
PR 03-SEP-1999; 99DE-01042095.  
PR 03-SEP-1999; 99DE-01042124.  
PR 03-SEP-1999; 99DE-01042129.  
PR 09-MAR-2000; 2000US-0187970P.  
XX  
XX (BADI ) BASF AG.  
XX  
XX Pompeius M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
XX WPI; 2001-137957/14.  
XX N-ESDB; AAF71779.  
XX  
XX Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway  
PT proteins, useful for producing fine chemicals in microorganisms,  
PT including organic acids, nonproteinogenic amino acids, and purine and  
PT pyrimidine bases.  
XX  
XX Claim 20; Page 234-235; 1737pp; English.  
XX  
XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic  
CC pathway (MP) proteins given in AAF79634 to AAF80211. The C. glutamicum MP  
CC nucleic acids are useful for the production of fine chemicals in  
CC microorganisms, including organic acids, nonproteinogenic amino acids,  
CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated  
CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,  
CC vitamins, cofactors, polyketides and enzymes  
XX  
XX Sequence 233 AA;

Query Match 98.9%; Score 1178; DB 4; Length 233;  
Best Local Similarity 100.0%; Pred. No. 1.4e-126;  
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MEIPIITGLLGASLLSIPQNVIVIKQIKREGIAVLVCLISDVFLFIAGTGLV 63  
DB 1 MEIPIITGLLGASLLSIPQNVIVIKQIKREGIAVLVCLISDVFLFIAGTGLV 60

QY 64 SNAAPIVLDIMRWGGIAYLLWFAVMAAKDANTNKVEAPQIIEETPTVDDTPLGGS 123  
DB 61 SNAAPIVLDIMRWGGIAYLLWFAVMAAKDANTNKVEAPQIIEETPTVDDTPLGGS 120

QY 124 TDTNRVREVSVDKQVWVKPMLMAVLITMLNPNAYLDAFVFIGVGAQYGDGTGRW 183  
DB 121 TDTNRVREVSVDKQVWVKPMLMAVLITMLNPNAYLDAFVFIGVGAQYGDGTGRW 180

QY 184 AGAFASLIWFLVGFGAALSRPLSPKVRWVNVVAVVMTALAKMLMG 236

Db 181 AGAFAASLIWFLVGFCAALSRPLSPKVRWVWVAVVMTALAIAKMLMG 233

RESULT 5  
AAU71888  
ID AAU71888 standard; protein; 233 AA.  
AC AAU71888;  
XX  
XX  
XX 26-FEB-2002 (first entry)  
DT  
DE C. glutamicum metabolic pathway protein encoded by gene #23.  
XX  
XX  
KW Metabolic pathway protein; MP; lysine biosynthesis pathway;  
KW methionine biosynthesis pathway; large-scale production of fine chemical;  
KW Corynebacterium diphtheriae; diphtheria.  
XX  
XX  
OS Corynebacterium glutamicum.  
XX  
XX WO2001:66573-A2.  
PN  
XX  
XX 13-SEP-2001.  
PD  
XX  
XX 22-DEC-2000; 2000WO-IB002035.  
PF  
XX  
XX 09-MAR-2000; 2000US-0187970P.  
PR  
XX 23-JUN-2000; 2000US-00606740.  
XX  
XX (BADI ) BASF AG.  
PA  
XX  
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
PI Kim J, Lee H, Hwang B;  
PI  
XX  
XX WPI; 2001-582369/65.  
DR N-PSDB; AAS96098.  
XX  
XX Nucleic acids encoding metabolic pathway proteins from Corynebacterium  
PT glutamicum, useful for producing methionine and lysine in Corynebacterium  
PT and Brevibacterium.  
PT  
XX  
XX Disclosure; Page 215-216; 316pp; English.  
PS  
XX  
XX The present invention relates to the isolation of novel Corynebacterium  
CC glutamicum genes (AAS96073-AAS96132) encoding metabolic pathway (MP)  
CC proteins. The metabolic pathway proteins of the invention include enzymes  
CC involved in the lysine and methionine biosynthetic pathways. The  
CC polynucleotide sequences of the invention can be used for the large-scale  
CC production and/or modulation of expression of fine chemicals such as  
CC lysine and methionine. The sequences of the invention may be used to  
CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a  
CC subject to detect diphtheria. AAU71863-AAU71922 represent the novel C.  
CC glutamicum metabolic pathway proteins of the invention  
XX  
SQ Sequence 233 AA;

Query Match 98.9%; Score 1178; DB 4; Length 233;  
Best Local Similarity 100.0%; Pred. No. 1.4e-126;  
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MEIITGILLGASLLSISGPNVLIKQIKREGILAVLVCLISDVFLFIAGTGLVDLL 63  
DB 1 MEIITGILLGASLLSISGPNVLIKQIKREGILAVLVCLISDVFLFIAGTGLVDLL 60  
QY 64 SNAAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIETPTVPDDTPLGGSAVA 123  
DB 61 SNAAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIETPTVPDDTPLGGSAVA 120  
QY 124 TDTNRVRVEVSDQRVWKPMMLAIVLTWLNPNAYLDAFVFIGVGAQYGDGRWIFA 183  
DB 121 TDTNRVRVEVSDQRVWKPMMLAIVLTWLNPNAYLDAFVFIGVGAQYGDGRWIFA 180  
QY 184 AGAFAASLIWFLVGFCAALSRPLSPKVRWVWVAVVMTALAIAKMLMG 236  
DB 181 AGAFAASLIWFLVGFCAALSRPLSPKVRWVWVAVVMTALAIAKMLMG 233

Db 181 AGAFAASLIWFLVGFCAALSRPLSPKVRWVWVAVVMTALAIAKMLMG 233

RESULT 6  
ADB66197  
ID ADB66197 standard; protein; 233 AA.  
XX  
XX  
XX ADB66197;  
AC  
XX  
XX 04-DEC-2003 (first entry)  
DT  
DE Protein encoded by C. glutamicum LysE gene.  
XX  
XX  
KW L-arginine production; coryneform bacteria; lysE; arginine repressor;  
KW argR; liver function promoting agent; amino acid infusion;  
KW amino acid pharmaceutical.  
XX  
XX  
OS Corynebacterium glutamicum.  
XX  
XX US2003113899-A1.  
PN  
XX  
XX 19-JUN-2003.  
PD  
XX  
XX 17-JUL-2002; 2002US-00196232.  
PF  
XX  
XX 25-JUL-2001; 2001JP-00224586.  
PR  
XX  
XX (AJIN ) AJINOMOTO CO INC.  
PA  
XX  
XX Yamaguchi M, Ito H, Gunji Y, Yasueda H;  
PI  
XX  
XX WPI; 2003-708853/67.  
DR N-PSDB; ADB66196.  
XX  
XX A microorganism comprising enhanced expression of the lysE gene is useful  
PT for enhanced production of L-arginine.  
PT  
XX  
XX Disclosure; Page 27-28; 36pp; English.  
PS  
XX  
XX The present invention relates to a method for producing L-arginine in a  
CC microorganism (e.g. coryneform bacteria) that has L-arginine producing  
CC ability and has been modified for enhanced expression of the lysE gene.  
CC The microorganism is also modified so that an arginine repressor (argR)  
CC does not function normally. The method of the invention is useful for the  
CC enhanced production of L-arginine which is useful in liver function  
CC promoting agents, amino acid infusion and comprehensive amino acid  
CC pharmaceuticals. The present sequence represents a protein encoded by the  
CC Corynebacterium glutamicum LysE gene. Note: The present sequence is given  
CC as SEQ ID No:25 in the Sequence Listing but is referred to as SEQ ID  
CC No:26 in the rest of the specification.  
XX  
SQ Sequence 233 AA;

Query Match 98.9%; Score 1178; DB 7; Length 233;  
Best Local Similarity 100.0%; Pred. No. 1.4e-126;  
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MEIITGILLGASLLSISGPNVLIKQIKREGILAVLVCLISDVFLFIAGTGLVDLL 63  
DB 1 MEIITGILLGASLLSISGPNVLIKQIKREGILAVLVCLISDVFLFIAGTGLVDLL 60  
QY 64 SNAAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIETPTVPDDTPLGGSAVA 123  
DB 61 SNAAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIETPTVPDDTPLGGSAVA 120  
QY 124 TDTNRVRVEVSDQRVWKPMMLAIVLTWLNPNAYLDAFVFIGVGAQYGDGRWIFA 183  
DB 121 TDTNRVRVEVSDQRVWKPMMLAIVLTWLNPNAYLDAFVFIGVGAQYGDGRWIFA 180  
QY 184 AGAFAASLIWFLVGFCAALSRPLSPKVRWVWVAVVMTALAIAKMLMG 236  
DB 181 AGAFAASLIWFLVGFCAALSRPLSPKVRWVWVAVVMTALAIAKMLMG 233

```
RESULT 7
AAG64047
ID AAG64047 standard; protein; 231 AA.
XX
AC AAG64047;
XX
DT 11-SEP-2001 (first entry)
XX
DE Corynebacterium thermoaminogenes lysin biosynthetic enzyme lysE.
XX
KW Heat-resistant; lysin biosynthesis; enzyme; coryneform;
KW aspartate-semialdehyde dehydrogenase; lysE.
XX
OS Corynebacterium thermoaminogenes.
XX
PN JP2001120270-A.
XX
PD 08-MAY-2001.
XX
PF 01-NOV-1999; 99JP-00311148.
XX
PR 01-NOV-1999; 99JP-00311148.
XX
PA (AJIN ) AJINOMOTO KK.
XX
DR WPI; 2001-364760/38.
DR N-PSDB; AAH45375.
XX
XX A heat-resistant lysin biosynthetic system enzyme gene of a high
XX temperature-resistant coryneform microbe.
XX
PS Claim 7; Page 24; 27pp; Japanese.
XX
CC The invention relates to a gene from a high temperature-resistant
CC coryneform microbe that encodes a heat-resistant lysin biosynthetic
CC enzyme. The enzyme has aspartate-semialdehyde dehydrogenase activity and
CC can be used for growing amino acid-producing microbes. The present amino
CC acid sequence corresponds to an enzyme of the invention
XX
SQ Sequence 231 AA;
Query Match 72.6%; Score 864.5; DB 4; Length 231;
Best Local Similarity 71.2%; Pred. No. 1.4e-90;
Matches 166; Conservative 30; Mismatches 32; Indels 5; Gaps 2;
Qy 4 MEIPIITGLLGASLLISIGPQNVIVIKQIKREGIIVLVCLISDVFLFIAGTLGV 63
Db 4 MEIPIITGLLGASLLISIGPQNVIVIKQIKREGIIVLVCLISDVFLFIAGTLGV 63
Qy 64 SNAAPIVLIDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIEETPTVDDTPLGSSAVA 123
Db 64 SDTAPIIDILRWCGIAYLLWFVMAAKDAMTNKVEAPQIIEETPTVDDTPLGSSAVA 118
Qy 124 TDTNRVVEVSDKQVVKPMMAIVLTWLNPNAYLDAPFVIGVGGAQYGDTRWIFA 183
Db 119 VTTQRPRLITSGTRQVVRPMMAIVLTWLNPNAYLDAPFVIGVGGAQYGDTRWIFA 178
Qy 184 AGAFAASLWFLVPGFGAALSRPLSPKVVWVWVAVVMTALAKMLMG 236
Db 179 AGAFAASLWFLVPGFGAALSRPLSPKVVWVWVAVVMTALAKMLMG 231
RESULT 8
ABR58214
ID ABR58214 standard; protein; 124 AA.
XX
AC ABR58214;
XX
DT 23-OCT-2003 (revised)
DT 11-AUG-2003 (first entry)
XX
DE LysE24 protein.
```

```
XX L-lysine; L-arginine; LysE24.
KW Corynebacterium glutamicum.
OS EPI266966-A2.
XX
PN 18-DEC-2002.
XX
PF 05-JUN-2002; 2002EP-00012539.
XX
PR 12-JUN-2001; 2001JP-00177075.
XX
PA (AJIN ) AJINOMOTO CO INC.
XX
PI Gunji Y, Yasueda H;
XX
DR WPI; 2003-241171/24.
DR N-PSDB; ACC80942.
XX
XX Novel DNA encoding variant of LysE protein from a coryneform bacterium,
XX when introduced into methanol assimilating bacterium, facilitates
XX excretion of L-lysine and/or L-arginine to outside of a cell.
XX
PS Claim 8; Page 20; 23pp; English.
XX
CC The present invention relates to DNA encoding variants of protein with
CC loop region and six hydrophobic helices which facilitates excretion of L-
CC lysine and/or L-arginine to outside of cell of a methanol assimilating
CC bacterium when introduced into the bacterium. The method is used for
CC encoding a protein which facilitates excretion of L-lysine, L-arginine or
CC both of these L-amino acids to outside of a cell of a methanol
CC assimilating bacterium when DNA of the method is introduced into the
CC bacterium. The present sequence represents a LysE24 protein from
CC Brevibacterium lactofermentum. (Updated on 23-OCT-2003 to standardise OS
CC field)
XX
SQ Sequence 124 AA;
Query Match 49.0%; Score 584; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 1e-58;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYMEIITGLLGASLLISIGPQNVIVIKQIKREGIIVLVCLISDVFLFIAGTLGV 60
Db 1 MYMEIITGLLGASLLISIGPQNVIVIKQIKREGIIVLVCLISDVFLFIAGTLGV 60
Qy 61 DLLSNAAPIVLIDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIEETPTVDDTPLG 118
Db 61 DLLSNAAPIVLIDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIEETPTVDDTPLG 118
RESULT 9
AAB01789
ID AAB01789 standard; protein; 211 AA.
XX
AC AAB01789;
XX
DT 03-JAN-2001 (first entry)
XX
DE Escherichia coli YggA amino acid excretion protein.
XX
KW E. coli; yggA gene; amino acid production; excretion protein gene;
KW amino acid excretion protein.
XX
OS Escherichia coli.
XX
PN EPI016710-A2.
XX
PD 05-JUL-2000.
XX
PF 17-DEC-1999; 99EP-00125263.
XX
```





XX PN WO9640893-A1.  
 XX PD 19-DEC-1996.  
 XX PF 06-JUN-1996; 96WO-US009122.  
 XX PR 07-JUN-1995; 95US-00487032.  
 XX PR 01-APR-1996; 96US-00630405.  
 XX PA (ASTR) ASTRA AB.  
 XX PI Smith D, Berglindh OT, Mellgaerd BL;  
 XX DR WPI: 1997-052306/05.  
 XX DR N-PSDB; AAT68221.  
 XX PT Helicobacter pylori nucleic acid sequences and related polypeptide(s) -  
 PT useful for vaccines to treat or prevent H. pylori infection, and to  
 PT detect Helicobacter.  
 XX PS Claim 73; Page 1359; 1481pp; English.  
 CC This sequence represents a H. pylori protein likely to contain five  
 CC membrane spanning regions. The protein may be used in a vaccine to  
 CC prevent or treat H. pylori infection or to identify H. pylori polypeptide  
 CC binding compounds, useful as potential H. pylori life cycle activators or  
 CC inhibitors. The genomic sequence of H. pylori (ATCC 55679) was determined  
 CC from overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides, and  
 CC the predicted coding regions defined by computer evaluation. To identify  
 CC likely H. pylori antigens for vaccine development, the amino acid  
 CC sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts  
 XX Sequence 215 AA;

Query Match 14.3%; Score 170.5; DB 2; Length 215;  
 Best Local Similarity 21.5%; Pred. No. 6.7e-11;  
 Matches 50; Conservative 48; Mismatches 92; Indels 43; Gaps 5;  
 QY 6 IFITGLLGASLLSIGPQNVIVKQIKREGLIAVLVCLISDVFLFIAGTGLVDLLN 65  
 Db 9 VFIEGGLAISLCAVAGQSLFIVERGMARNYVFLICALCFMCDIVLMSMGVFGVGNYP- 67  
 QY 66 AAPIVLDI-MRWGGIAYLLNFAVMAAKDAMTKVEAPQIIETEPTVPDDTPLGGSVAT 124  
 Db 68 AKNLYLSLFLNLFQAVFTGFYALALK-----T 95  
 QY 125 DTRNRVREVSVDKQVWVKPMLMAIVLTWLNPNAYLDADFVFIGVGQYGDGTGRWIFAA 184  
 Db 96 LFQTFKKQVQTEPKLSLKKTLFTLGVTLLNPQVLYLEWFLIGASAMSNLQKVFVFLA 155  
 QY 185 GAFAASLWFPPL-----VGFGAALSRPLSSPKVWRWVNVAVVMTALAIAKL 232  
 Db 156 GTLSAASFLLICTMSLRG----SKLNNQKIFMGVNLVFTVNTGLTSLVTL 204

RESULT 14  
 ABM72206  
 ID ABM72206 standard; protein; 205 AA.  
 XX AC ABM72206;  
 XX DT 20-NOV-2003 (first entry)  
 XX DE Staphylococcus aureus protein #1446.  
 XX KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;  
 KW enzymatic assay; antibiotic target.

XX OS Staphylococcus aureus.  
 XX PN WO200294868-A2.  
 XX PD 28-NOV-2002.  
 XX PF 27-MAR-2002; 2002WO-IB002637.  
 XX PR 27-MAR-2001; 2001GB-00007661.  
 XX PA (CHIR-) CHIRON SPA.  
 XX PI Masignani V, Mora M, Scarselli M;  
 XX DR WPI: 2003-120786/11.  
 XX DR N-PSDB; ACF73766.  
 XX PT New Staphylococcus aureus protein, useful as a vaccine for treating or  
 PT preventing Staphylococcal infection, specifically an infection caused by  
 PT S. aureus, e.g. sepsis.  
 XX PS Claim 1; SEQ ID NO 2892; 49pp; English.  
 CC The invention relates to novel genes and encoded proteins from  
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a  
 CC nucleic acid encoding the protein, or an antibody to the protein, is  
 CC useful as a pharmaceutical, particularly as a vaccine for treating or  
 CC preventing infection due to Staphylococcus bacteria, specifically an  
 CC infection caused by S. aureus. The composition is particularly useful for  
 CC treating or preventing sepsis in a patient. The composition can also be  
 CC used for diagnostics. The protein is also used in an assay for enzymatic  
 CC studies and as a target for antibiotics. This sequence represents one of  
 CC the novel S. aureus proteins of the invention  
 XX Sequence 205 AA;

Query Match 13.8%; Score 164; DB 6; Length 205;  
 Best Local Similarity 21.4%; Pred. No. 3.5e-10;  
 Matches 49; Conservative 47; Mismatches 93; Indels 40; Gaps 4;  
 QY 8 ITGILLGASLLSIGPQNVIVKQI---KREGLIAVLVCLISDVFLFIAGTGLVDLLS 64  
 Db 6 IHGFIILGILPLGAQNVIFNQGANQPKYRYPALITAGLSLSLLIIIAVVGVSIII 65  
 QY 65 NAAPIVLDI-MRWGGIAYLLNFAVMAAKDAMTKVEAPQIIETEPTVPDDTPLGGSVAT 124  
 Db 66 MSLPVLQAIIVIGLIFLWYAW-----TIWEDKE-----ST 97  
 QY 125 DTRNRVREVSVDKQVWVKPMLMAIVLTWLNPNAYLDADFVFIGVGQYGDGTGRWIFAA 184  
 Db 98 DGETQIMSP-----MKQVSFALSLSLPHALITGIVIGSAAALYGSNKIAFTI 148  
 QY 185 GAFAASLWFPPLVGFGAALSRPLSSPKVWRWVNVAVVMTALAIAKL 233  
 Db 149 ACISVSWLWFLAILGKRWGSDIKTKLITINKISSIIIIIVLMLIL 197

RESULT 15  
 ABM67507  
 ID ABM67507 standard; protein; 205 AA.  
 XX AC ABM67507;  
 XX DT 20-NOV-2003 (first entry)  
 XX DE Photorhabdus luminescens protein sequence #604.  
 XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
 KW detection; food; gene expression; plant; animal; microorganism; toxin;  
 KW antibiotic; biopesticide; virulence factor; disease model; plague;  
 KW whooping cough.  
 XX

Search completed: March 12, 2004, 18:24:47  
Job time : 63 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2004, 18:24:51 ; Search time 22 Seconds  
(without alignments)  
553.806 Million cell updates/sec

Title: US-09-105-117K-2  
Perfect score: 1191  
Sequence: 1 MVMEIFITGILLGASLLLS.....INVVAVMTALAIKMLMG 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pap.\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pap.\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pap.\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pap.\*  
5: /cgn2\_6/prodata/2/iaa/PCJUS COMB.pap.\*  
6: /cgn2\_6/prodata/2/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	341	28.6	213	4	US-09-489-039A-11845
2	332	27.9	222	4	US-09-252-991A-20747
3	310	26.0	214	4	US-09-543-681A-4972
4	298	25.0	211	4	US-09-328-352-5737
5	138.5	11.6	212	4	US-09-543-681A-4767
6	135.5	11.4	260	4	US-09-252-991A-25992
7	121.5	10.2	235	4	US-09-328-352-8249
8	119.5	10.0	173	4	US-09-134-001C-4026
9	119.5	10.0	222	4	US-09-328-352-6471
10	118.5	9.9	250	4	US-09-543-681A-4487
11	117.5	9.7	219	4	US-09-252-991A-18746
12	115.5	9.7	216	4	US-09-489-039A-7731
13	113.5	9.5	228	4	US-09-543-681A-4854
14	111	9.3	222	4	US-09-252-991A-26644
15	108.5	9.1	238	4	US-09-328-352-7122
16	108	9.1	217	4	US-09-489-039A-8076
17	104.5	8.8	211	4	US-09-543-681A-6708
18	104	8.7	249	4	US-09-252-991A-25717
19	101.5	8.5	220	4	US-09-328-352-4903
20	99	8.3	208	4	US-09-328-352-5111
21	98.5	8.3	1107	4	US-09-489-039A-8890
22	96	8.1	1421	4	US-09-252-991A-17805
23	94.5	7.9	240	4	US-09-328-352-7537
24	94	7.9	211	4	US-09-489-039A-9608
25	94	7.9	552	4	US-09-489-039A-7419
26	93	7.8	228	4	US-09-252-991A-23837
27	91.5	7.7	484	4	US-09-634-238-216

28	91	7.6	211	4	US-09-252-991A-20747	Sequence 20747, A
29	90	7.6	205	4	US-09-489-039A-12661	Sequence 12661, A
30	89.5	7.5	205	4	US-09-396-357-2	Sequence 2, Appli
31	88.5	7.4	350	4	US-09-252-991A-32358	Sequence 32358, A
32	87	7.3	398	4	US-09-328-352-6411	Sequence 6411, Ap
33	87	7.3	451	4	US-09-107-532A-5352	Sequence 5352, Ap
34	86.5	7.3	226	4	US-09-252-991A-31610	Sequence 31610, A
35	86.5	7.3	456	4	US-09-543-681A-6928	Sequence 6928, Ap
36	86	7.2	485	4	US-09-134-001C-6031	Sequence 6031, Ap
37	85.5	7.2	224	4	US-09-134-001C-4608	Sequence 4608, Ap
38	85.5	7.2	269	4	US-09-489-039A-7320	Sequence 7320, Ap
39	85.5	7.2	472	4	US-09-252-991A-18544	Sequence 18544, A
40	85.5	7.2	475	4	US-09-252-991A-29697	Sequence 29697, A
41	84.5	7.1	384	4	US-09-252-991A-19333	Sequence 19333, A
42	84	7.1	214	4	US-09-328-352-6695	Sequence 6695, Ap
43	84	7.1	241	4	US-09-328-352-7159	Sequence 7159, Ap
44	84	7.1	289	4	US-09-328-352-6229	Sequence 6229, Ap
45	83.5	7.0	210	4	US-09-328-352-4775	Sequence 4775, Ap

## ALIGNMENTS

RESULT 1  
US-09-489-039A-11845  
; Sequence 11845, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT FILING DATE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 11845  
; LENGTH: 213  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-11845

Query Match 28.6%; Score 341; DB 4; Length 213;  
Best Local Similarity 32.5%; Pred. No. 8.3e-32;  
Matches 77; Conservative 43; Mismatches 83; Indels 34; Gaps 4;  
QY 1 MVMEIFITGILLGASLLLSIGPQNVIVKQIKREGIAVLVCLISDVFLFIAGTIGV 60  
DB 1 ILMFTYFQGLGAAMILPLGPNQAFVQIRQYHIMIALLCVSDLLICAGIFGG 60  
QY 61 DLLSNAPIVLDIMRWGTAIVLLNFVAVMAKADMTKNKVEAPQITEPTVDDTPLGGS 120  
DB 61 SALLMQSPWLLAVTWGGVAFLLWYFGCALKTAFPSQLEL-----100  
QY 121 AVATDTRNREVEVSVDKQKRVKPKMLMAIVLTWLNPNAYLDADFVFGVGAQYG-DYGR 179  
DB 101 -----ANAEVQQGRW-KIIITMLAVTWLNPHVYLDTFVLGSLGSGQLAVEPKR 148  
QY 180 WFAAGAAFAASLIWFFVFGGAALSRLSPKVKWVNVVAVMTALAIKMLMG 236  
DB 149 W-FALGTISNLFMFFGLALLAALAPRLTARAQRINIVVGAVMFWFIQALREG 204

RESULT 2  
US-09-252-991A-24479  
; Sequence 24479, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136

[illegible]

```

Query Match      10.2%; Score 121.5; DB 4; Length 235;
Best Local Similarity 23.0%; Pred. No. 4.8e-06;
Matches 57; Conservative 41; Mismatches 89; Indels 61; Gaps 12;

QY      3 IMEITGLLGLGASLLSIGPQNVIVIKQGIKREGILAVLLV---CLISDVLFITAGTLG 59
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      29 LQIIALFGLIC-LAWLITGPPNMILISRSIS-QGKIAGFISLGGVAGVFVYMLCASFG 86
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      60 VDLLSNAAPIVLDIRMGGGIAYLLWFPAVMAAKADMTNKVEAPQIIEETEPTVDDTPLGG 119
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      87 ITALVAVPYAVDTIRIAGAMYLWLAWKALR-----PNAAPIFN 126
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      120 SAVATDENRVVEVSVDQRVWVKPMLMAIVUTWLNPN-----AYLDAFVP--IGGVG 171
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      127 VK-----DLAVDSP---LKLFLMGFLTNLNPKIAMIYLSLLPQIHPOQSGSIL 172
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      172 AOVGDTGR-WIFAAGAFASLIWPELPFGCAALSRPLSSPKWV---RWINVVAVVMT 226
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      173 AQSIQLGTIQTIVSVSVNA-----LIVFANGSIALFLQKPELWASIQRW---VMGTVLA 223
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      227 ALAIKML 234
      : : : : :
Db      224 GLAVRILL 231
      : : : : :

```

```

1 / GENERAL INFORMATION:
2 / APPLICANT: Lynn Doucette-Stamm et al
3 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
4 / TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
5 / FILE REFERENCE: GTC-007
6 / CURRENT APPLICATION NUMBER: US/09/134,001C
7 / CURRENT FILING DATE: 1998-08-13
8 / PRIOR APPLICATION NUMBER: US 60/064,964
9 / PRIOR FILING DATE: 1997-11-08
10 / PRIOR APPLICATION NUMBER: US 60/055,779
11 / PRIOR FILING DATE: 1997-08-14
12 / NUMBER OF SEQ ID NOS: 5674
13 / SEQ ID NO 4026
14 / LENGTH: 173
15 / TYPE: PRT
16 / ORGANISM: Staphylococcus epidermidis
17 / US-09-134-001C-4026

```

```

Query Match      10.0%; Score 119.5; DB 4; Length 173;
Best Local Similarity 23.0%; Pred. No. 5.3e-06;
Matches 46; Conservative 32; Mismatches 83; Indels 39; Gaps 4

QY 39 IAVTLVCLISDVFLFIAGTILGVDTLLSNAAPIVLIDIMRWGGIAYLLMFVFAAKQAMTKV 98
      : : : : | | | : | | : : : : : : : : : : : : : : : : : : : : : :
Db 8 LPVITTAGLCDTFLIVAILGVSLLISMTPTQLFYITIGFLMTN-----ANSLMT 60
      : : : : | | | : | | : : : : : : : : : : : : : : : : : : : : : :
QY 99 EAPQIIEETPTVPDPTPLGGSAVATDTRNRVRVSVSDQRVWVKPMLMAIVLTMLNPN 158
      : : : : | | | : | | : : : : : : : : : : : : : : : : : : : : : :
Db 61 EKPNSNIEIEP-----MSKKQ-----ILFALSLSLLNPH 90
      : : : : | | | : | | : : : : : : : : : : : : : : : : : : : : : :
QY 159 AYLDAPFVIGGVAQYGTGRWIFPAAGAAFLIWFPLVDFGAAALSRPLSPKVKRWIN 218
      : : : : | | | : | | : : : : : : : : : : : : : : : : : : : : : :
Db 91 AIMDTGVGIGTSASVYDGVKVFSLATISVSWIWFVFLAILGRITGKIDKSGKYIVILN 150
      : : : : | | | : | | : : : : : : : : : : : : : : : : : : : : : :
QY 219 VV--VAVVTALAIKMLMG 236
      : : : : | | | : | | : : : : : : : : : : : : : : : : : : : : : :
Db 151 KVSSTVIVIGVLIILKNIVG 170
      : : : : | | | : | | : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-09-328-352-6471
; Sequence 6471, Application US/09328352
Patent No. 6562958

```

```
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-039A
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6471
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-6471

Query Match      10.8%; Score 119.5; DB 4; Length 222;
Best Local Similarity 23.1%; Pred. No. 7.7e-06;
Matches 54; Conservative 38; Mismatches 95; Indels 47; Gaps 9;

QY 6 IFITGLLGASLLSIGPQNVLVKQIKREGLIALLVCLISDVFLFIAGTGVDDLNS 65
Db 24 IFIAFWSVGLFIITPGADWAYAISAGIKGVVVPVAVAGMLFGHEFITILLVAGVGLVA 83
QY 66 AAPIVLDIMRWGGIAYLLWFAVMAAKDAMTKVEAPQIIEETPTVDDTPLGSSAVATD 125
Db 84 NNPTALMILTAVAGSAYLLWMI-----NLLTLP-----PT-PKKS-----GS----- 119
QY 126 TRNRVRVSVDKQVWVKPMLMAIVLTWLPNAYLDADFVIGVGAGYQDGTGRW----- 180
Db 120 -----EKAQSWLRWATKGVVVSGLNPKVFL---LFLALLPQFIDTTASMSVTTQ 165
QY 181 IFAGA--PAASLWFLVFGGAAL--SRPLSPKVKRWVNVVAVMTALAI 230
Db 166 ILAFGVVHIISCAIYLMVYGSEALKTRPQAALVGRFSGGLMVVATCLLI 219

RESULT 10
US-09-543-681A-4487
; Sequence 4487, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1998-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4487
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-4487

Query Match      9.9%; Score 118.5; DB 4; Length 250;
Best Local Similarity 21.9%; Pred. No. 1.2e-05;
Matches 52; Conservative 35; Mismatches 101; Indels 49; Gaps 8;

QY 6 IFITGLLGASLLSIGPQNVLYI---KQIKREGLIALLVCLISDVFLFIAGTGLVD 61
Db 53 IFIT-----LVPGENSIFVETSSAKGVK--GGYKAALGVFTGDALLIFIAFLGVA 101
QY 62 LLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTKVEAPQIIEETPTVDDTPLGSSA 121
Db 102 SLVNTSPVFVIKYGALYLYTLGLKLYATFHKKKEQPE----- 142
QY 122 VATDTRNRVRVSVDKQVWVKPMLMAIVLTWLPNAYL--DAFVIGVGAGYQDGTGRW 180
Db 143 -----QVAVSVKAKNGLYIK-----ALFLSLNPKMIIFFVSFFIQFIDPKYENAGVP 191
QY 181 IFAGAF--AASLWFLVFGGAALSRLS--SPKVRWNVNVVAVMTALAIKMLI 234
```

```
Db 192 PFVLGVLEICSMLYSLVLIFFGVGVAITNKVKHNRKRLASLNSCIGAVFLFGAKIAL 248

RESULT 11
US-09-252-991A-18746
; Sequence 18746, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18746
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-18746

Query Match      9.9%; Score 117.5; DB 4; Length 206;
Best Local Similarity 25.4%; Pred. No. 1.2e-05;
Matches 64; Conservative 31; Mismatches 82; Indels 75; Gaps 14;

QY 4 MEIFITGLLGASLLSIGPQNVLVKQIKREGLIA-----VLVCLISDVFLFIAGTGLG 59
Db 10 MQSLVFPFLFAVVASITPGPTNLTLSNS--QRHGLAAAPVILGCAVAAALILLILG-LG 67
QY 60 VDLISNAAPVLDIMRW---GGIAYLLWFAVMAAKDAMTKVEAPQIIEETPTVDDTPT 116
Db 68 LGELLRHPPLQQGLAMLGWGLSYLANSLSRSAGG-----IDGAE-----P 110
QY 117 -----LGGSAVATDTRNRVRVSVDKQVWVKPMLMAIVLTWLPNAY---LDAFVFIG 168
Db 111 RRLGVLGAA-----LQLVNPKAMMALAALFA 140
QY 169 GVGAGYQDGTGR--WIFAGAFASLI---LWPLVFGGAALSRLSPKVKRWNVNVVAV 223
Db 141 GEGA--GQAGRIGLLALLFFLVSLPCLASW--ALLGVGSARL---LRSFLMKRFRNQGMAL 194
QY 224 VMTALAIKMLM 235
Db 195 LLLASAWAALL 206

RESULT 12
US-09-489-039A-7731
; Sequence 7731, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7731
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-7731

Query Match      9.7%; Score 115.5; DB 4; Length 219;
Best Local Similarity 24.4%; Pred. No. 2.2e-05;
Matches 50; Conservative 37; Mismatches 99; Indels 19; Gaps 8;
```

```

QY 7 FITGLGASLLLSIGPQNVLVKQIK---REGLIAVLVCLISVFLFIAGTGLGVDLL 63
Db 17 LITLVGAVFILLVPGPTFFVLKATGIAHGKKGKYLAAAGV-FIGDAVLMFLAFAGVATL 75
QY 64 SNAAPVLDIMRWGGIAYILLW-----FVMAAKDAMTN-KVEAPQIIEETEPTVPDDTP 116
Db 76 IKTTPELVFNVRVYLGAIYLLMLGGKMLYAVLTQRDQSDASAEPAASAILKSLTSLTNP 135
QY 117 LGGSVATDTRNRVRVEVDKQVWVKPMLMAIVLTWLPNAYLDADFVFTGGVGAQYGD 176
Db 136 ---KAILFVVSFPQF-IDVNKATPGVAFILALTLEVIS-FCYMSFLLSGSEFVIRYK 190
QY 177 TGRWIFAAAGFAASLIMFPLVFGCA 201
Db 191 TRKIAKLG---NSLILGVFVGFAA 212

RESULT 13
US-09-543-681A-4854
; Sequence 4854, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543, 681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4854
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4854

Query Match 9.5%; Score 113.5; DB 4; Length 228;
Best Local Similarity 23.1%; Pred. No. 4e-05;
Matches 56; Conservative 36; Mismatches 87; Indels 63; Gaps 10;

QY 7 FITGLGASLLLSIGPQNVLVKQIKR---EGLIAVLVCLISVFLFIAGTGLGVDLL 63
Db 29 YLAGMFP---IIIVGPNLSVLKTSRSGVGAGYRAALGV-FIGDAILFLAFIGVASV 84
QY 64 SNAAPVLDIMRWGGIAYILLWFAV-----MAAKDAMTNKVEAPQIIEETEPTVPDDTP 116
Db 85 IKASPLVFTVIRFLGAFYLLYLGKTIHANFFAKKQSHT----- 123
QY 117 LGGSVATDTRNRVRVEVDKQVWVKPMLMAIVLTWLPNNA---YLDADFVFTGGVGAQ 173
Db 124 -----ETNCIOEVRPK-----ALTSLTPKRALFYSFPVQF-IDFN 161
QY 174 YGDTG--RWIFAGFAASLIMFPLVFGGAALSRPLSSPK-VWRWNVVAVVMTALAI 230
Db 162 YAHGLSYLILASMLEAFSFLIFGVALARFFGSRKNTAKLNGIVGLFPMGFAT 221
QY 231 KL 232
Db 222 KL 223

RESULT 14
US-09-252-991A-26644
; Sequence 26644, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```

```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26644
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26644

Query Match 9.3%; Score 111; DB 4; Length 222;
Best Local Similarity 22.6%; Pred. No. 7.6e-05;
Matches 53; Conservative 27; Mismatches 87; Indels 68; Gaps 8;

QY 11 LLLGASLLLSIGPQNVLVKQIGI---KREGLIAVLVCLISVFLFIAGTGLGVLLSNA 67
Db 19 LIALAVVLYVPGPDMELLFQTGARQGRRAALVTALGLALARACHVLMMAAT-GIALLPETA 77
QY 68 PIVLDIMRWGGIAYILLWFAVMAAKDAMTNKVEAPQIIEETEPTVPDDTPLGGSVATDTR 127
Db 78 PWTFLVRLGAAAYLAWLGLQMLRG-----GGIALPTSDA 112
QY 128 NRVRVEVDKQVWVKPMLMAIVLTWLPNAYLDADFVFI-----GVGGAQY----- 174
Db 113 GSAPVPHADR-----RALLRGLTLLNLPKALLPCSVLLPQFVSPGAGSLAVQPAALGT 167
QY 175 -----GDTORWIFAAAGFAASLIW---FPLVFGCA-AALSRPL 208
Db 168 VLVLVGLAFCAVALAGGRGLRWLASPRARLQWQGGILLIGFVRLALLRQL 222

RESULT 15
US-09-328-352-7122
; Sequence 7122, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7122
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7122

Query Match 9.1%; Score 108.5; DB 4; Length 238;
Best Local Similarity 24.3%; Pred. No. 0.00016;
Matches 53; Conservative 26; Mismatches 82; Indels 57; Gaps 9;

QY 22 GPQN--VLVTKQIKREGLIAVLVCLISVFLFIAGTGLGVLLSNAAPVLDIMRWGGI 79
Db 56 GTFNFIILSSHYKISKYLPVILGSCIGNALLVWVGIGLGTILAYFVIOKIMWSG- 114
QY 80 AYLLWFAVMAAKDAMTNKVEAPQIIEETEPTVPDDTPLGGSVATDTRNRVRVEVDKQ 139
Db 115 --LIWLVTLWANKLYN---FVLSLEKNEQYF---PIGPKA----- 147
QY 140 RWVKPMLMAIVLTWLPNAYLDADFVIGVGAQYGDTR-----WTEAAGAFAA 189
Db 148 -----AFILMQAINPKTWMAFAVI---SVYTKQGDILVNVVLSLSCIFLLIAPPC 194
QY 190 SLIWFPLVFGGAALSRPLSSPKVWRWNVVAVVMTA 227
Db 195 LYLW-ALVG---RLSTRLLSKPKHINIFKNKIAIILLA 228

Search completed: March 12, 2004, 18:27:23

```

Thu Mar 18 12:34:41 2004

us-09-105-117k-2.ra1

Page 6

Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: March 18, 2004, 06:36:47 ; Search time 43 Seconds  
(without alignments)  
1421.243 Million cell updates/sec

Title: US-09-105-117K-2  
Perfect score: 1191  
Sequence: 1 MYMEIFITGLLGASLLS.....INWVAVMTALIKMLMG 236

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25895339 residues  
Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	1191	100.0	236	9	US-09-738-626-6955
2	1191	100.0	236	14	US-10-166-142-8
3	1178	98.9	233	10	US-09-746-6608-52
4	1178	98.9	233	14	US-10-196-232-25
5	584	49.0	123	14	US-10-166-142-10
6	301	25.3	203	14	US-10-166-142-10
7	133	11.1	224	14	US-10-156-761-7796
8	132.5	11.1	226	9	US-09-738-626-3665
9	128.5	10.8	212	12	US-10-374-9038-6
10	119	10.0	620	12	US-10-282-122A-77618
11	109.5	9.2	210	14	US-10-156-761-14994
12	103.5	8.7	181	12	US-10-282-122A-44873
13	98.5	8.3	217	14	US-10-156-761-10918
14	95	8.0	295	12	US-10-282-122A-77782
15	93.5	7.9	452	9	US-09-815-242-10183

16	93.5	7.9	452	12	US-10-282-122A-56572
17	93	7.8	211	12	US-10-282-122A-47826
18	93	7.8	459	9	US-09-738-626-6998
19	92.5	7.8	298	14	US-10-156-761-8730
20	91.5	7.7	446	9	US-09-951-780-2
21	91.5	7.7	446	9	US-09-738-626-6356
22	91.5	7.7	446	12	US-10-627-476-426
23	91.5	7.7	484	15	US-10-264-213-124
24	91.5	7.7	485	14	US-10-080-170-297
25	91	7.6	252	12	US-10-425-114-65374
26	91	7.6	361	14	US-10-156-761-9054
27	91	7.6	390	15	US-10-369-493-8216
28	91	7.6	624	9	US-09-738-626-4289
29	91	7.6	628	12	US-10-627-476-550
30	91	7.6	628	12	US-10-627-476-666
31	90.5	7.6	472	14	US-10-080-170-380
32	89.5	7.5	205	9	US-09-927-395-2
33	89.5	7.5	205	9	US-09-847-392-2
34	89.5	7.5	340	9	US-09-815-242-11980
35	89	7.5	477	15	US-10-369-493-18633
36	88.5	7.4	436	14	US-10-081-872-28
37	88.5	7.4	436	15	US-10-385-305-28
38	88	7.4	436	14	US-10-081-872-42
39	88	7.4	436	14	US-10-228-063-1
40	88	7.4	436	15	US-10-385-305-42
41	88	7.4	454	14	US-10-228-063-13
42	88	7.4	460	14	US-10-228-063-14
43	88	7.4	460	14	US-10-228-063-33
44	88	7.4	460	14	US-10-228-063-35
45	88	7.4	518	14	US-10-228-063-15

## ALIGNMENTS

### RESULT 1

US-09-738-626-6955  
; Sequence 6955, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 6955  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6955

Query Match 100.0%; Score 1191; DB 9; Length 236;  
Best Local Similarity 100.0%; Pred. No. 6.1e-117;  
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MYMEIFITGLLGASLLSIGPQNVILVKGKIKREGILAVLVCLISDVFLIAGTLGV 60

```

; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 52
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-746-660A-52

Query Match      98.9%; Score 1178; DB 10; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.4e-115;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 MEIFITGLLGASLLISGPNVLVVKQIKREGIAVLVCLISDVFLFIAGTLGV 63
Db      1 MEIFITGLLGASLLISGPNVLVVKQIKREGIAVLVCLISDVFLFIAGTLGV 60

QY      64 SNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIETEPTVDDTPLGSSAVA 123
Db      61 SNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIETEPTVDDTPLGSSAVA 120

QY      124 TTRNRVRVSVSKORVWVKPMLMAVLTLNPNAYLDAFVFIGVGGAQYGDTRWIFA 183
Db      121 TTRNRVRVSVSKORVWVKPMLMAVLTLNPNAYLDAFVFIGVGGAQYGDTRWIFA 180

QY      184 AGAFAASLIWFLVGVGGAALSRPLSSPKVWRWVAVVMTALAIKMLMG 236
Db      181 AGAFAASLIWFLVGVGGAALSRPLSSPKVWRWVAVVMTALAIKMLMG 233

RESULT 4
US-10-196-232-25
; Sequence 25, Application US/10196232
; Publication No. US20030113899A1
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, MIKIKO
; APPLICANT: ITO, HISAO
; APPLICANT: GUNJI, YOSHIYA
; APPLICANT: YASUEDA, HISASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: 225391USO
; CURRENT APPLICATION NUMBER: US/10/196,232
; CURRENT FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: JP 2001-224586
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-10-196-232-25

Query Match      98.9%; Score 1178; DB 14; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.4e-115;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; TITLE OF INVENTION: CORNYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: 223789US
; CURRENT APPLICATION NUMBER: US/10/166,142
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: JP 2001-1777075
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
; US-10-166-142-8

Query Match      100.0%; Score 1191; DB 14; Length 236;
Best Local Similarity 100.0%; Pred. No. 6.1e-117;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MYMEIFITGLLGASLLISGPNVLVVKQIKREGIAVLVCLISDVFLFIAGTLGV 60
Db      1 MYMEIFITGLLGASLLISGPNVLVVKQIKREGIAVLVCLISDVFLFIAGTLGV 60

QY      61 DLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIETEPTVDDTPLGSS 120
Db      61 DLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIETEPTVDDTPLGSS 120

QY      121 AVATDTRNRVRVSVSKORVWVKPMLMAVLTLNPNAYLDAFVFIGVGGAQYGDTRW 180
Db      121 AVATDTRNRVRVSVSKORVWVKPMLMAVLTLNPNAYLDAFVFIGVGGAQYGDTRW 180

QY      181 IFAGAFASLIWFLVGVGGAALSRPLSSPKVWRWVAVVMTALAIKMLMG 236
Db      181 IFAGAFASLIWFLVGVGGAALSRPLSSPKVWRWVAVVMTALAIKMLMG 236

RESULT 3
US-09-746-660A-52
; Sequence 52, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORNYNEBACTERIUM GLUTAMICUM GENES ENCODING

```



QY 4 MEIFITGLLGASLLISIGPQNVLVKQIKREGIAVLVCLISDVFLFIAGTLGVDDL 63  
Db 1 MEIFITGLLGASLLISIGPQNVLVKQIKREGIAVLVCLISDVFLFIAGTLGVDDL 60  
QY 64 SNAPIVLDIMRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIETEPTVPDDTPGGSAVA 123  
Db 61 SNAPIVLDIMRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIETEPTVPDDTPGGSAVA 120  
QY 124 TDRNRVRVSVVDKQVWVKPMLMAIVLTWLPNAYLDARFVFGVGAQYGDGTGRWIFA 183  
Db 121 TDRNRVRVSVVDKQVWVKPMLMAIVLTWLPNAYLDARFVFGVGAQYGDGTGRWIFA 180  
QY 184 AGAFAASLIWFFLVGGAALSRPLSSPKVWRWVNVVAVMTALAIKMLMG 236  
Db 181 AGAFAASLIWFFLVGGAALSRPLSSPKVWRWVNVVAVMTALAIKMLMG 233

RESULT 5  
US-10-166-142-10  
; Sequence 10, Application US/10166142  
; Publication No. US20030124687A1  
; GENERAL INFORMATION:  
; APPLICANT: YASUEDA, HISASHI  
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL  
; FILE REFERENCE: 223789US  
; CURRENT APPLICATION NUMBER: US/10/166,142  
; PRIOR FILING DATE: 2002-06-11  
; PRIOR APPLICATION NUMBER: JP 2001-1777075  
; PRIOR FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 10  
; LENGTH: 124  
; TYPE: PRT  
; ORGANISM: Brevibacterium lactofermentum  
US-10-166-142-10

Query Match 49.0%; Score 584; DB 14; Length 124;  
Best Local Similarity 100.0%; Pred. No. 2e-53;  
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WYMEIFITGLLGASLLISIGPQNVLVKQIKREGIAVLVCLISDVFLFIAGTLGV 60  
Db 1 WYMEIFITGLLGASLLISIGPQNVLVKQIKREGIAVLVCLISDVFLFIAGTLGV 60  
QY 61 DLLSNAAPIVLDIMRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIETEPTVPDDTPGL 118  
Db 61 DLLSNAAPIVLDIMRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIETEPTVPDDTPGL 118

RESULT 6  
US-10-156-761-8659  
; Sequence 8659, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 8659  
; LENGTH: 203  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-8659

Query Match 25.3%; Score 301; DB 14; Length 203;  
Best Local Similarity 32.1%; Pred. No. 2.3e-23;  
Matches 72; Conservative 32; Mismatches 86; Indels 34; Gaps 3;

QY 10 GLLLGASLLISIGPQNVLVKQIKREGIAVLVCLISDVFLFIAGTLGVDDLSPAPI 69  
Db 11 GFGTGLSLIVAIGQNAFVLRQGIERRDAVLAIVGICALSALLIALGVGGVAVVWAF 70  
QY 70 VLDIMRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIETEPTVPDDTPGGSAVATDTR 129  
Db 71 ALTAVALYGGAFLLVYGALARRVF-----RPAURACAPESR 110  
QY 130 VRVEVSVDKQVWVKPMLMAIVLTWLPNAYLDARFVFGVGAQYGDGTGRWIFAAGAF 189  
Db 111 RRA-----VLTCLALTWLPNPHVVDVFLVGLSIAADRGSL-RWTEGLGAALA 156  
QY 190 SLIMEPLVGGGAALSRPLSSPKVWRWVNVVAVMTALAIKLM 233  
Db 157 SLCWFAALGFGSRLGRLGARPSAWRVLDAVVAATMLTWGATLI 200

RESULT 7  
US-10-156-761-7796  
; Sequence 7796, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 7796  
; LENGTH: 224  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-7796

Query Match 11.2%; Score 133; DB 14; Length 224;  
Best Local Similarity 24.2%; Pred. No. 1.3e-05;  
Matches 62; Conservative 40; Mismatches 92; Indels 62; Gaps 15;

QY 4 MEIFITGLLGASLLISIGPQNVLVKQIKREGIAVLVCLISDVFLFIAGTLGV 55  
Db 1 MSVDIVG-FIGVVLVAVVPGDFLVVRSATEHPAKGRAALGAQSGUCV-----HMLA 54  
QY 56 GTLVGDLISNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIETEPTVP 108  
Db 55 AAVGLSLIAARSPAVYDAIRLGAALVVLGVRAVLAARARARAGREAVGGVDEGTD 114  
QY 109 PTVP-DUTPVGSAVATDTRNRVRVSVVDKQVWVKPMLMAIVLTWLPNAYLDARF 167  
Db 115 PRTPPEEAPRGR-----WRSQGTQGTFLTNVLPNKAAL-----FF 149  
QY 168 GGVGAQY-----GDTGRWIFAAGAFASL-----IWFPLVGGGAALSRPLSPKV 215

Db 150 LSTLPQVHGGSSTQPIPLGTLDTLVIGVAYFALVAV-AARLRARFLARPKVRHGWEIT 208  
 QY 216 --WINVVAVMTALA 229  
 Db 209 TGLWLTATIGGVAAAA 224

RESULT 8  
 US-09-738-626-3665  
 ; Sequence 3665, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIALI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 3665  
 ; LENGTH: 226  
 ; TYPE: PRT  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-3665

Query Match 11.1%; Score 132.5; DB 9; Length 226;  
 Best Local Similarity 24.2%; Pred. No. 1.5e-05;  
 Matches 58; Conservative 45; Mismatches 98; Indels 39; Gaps 12;  
 QY 7 FTGGLGASLLSISGPQNVLIKQIK--REGLIAVLLVCLISDVFLIAGTLGVDLLS 64  
 Db 6 FLALFLVTAATSPGDLQIILSAKRRDGLVTAVGLWGNISWI-IALSLGLSALI 64  
 QY 65 NAAPVLDIMRWGGIAYLLMFVAVMAAKDMTKVKEAPQIIETEPTVDDTPLGGSAVAT 124  
 Db 65 STYPAILNLLQVLGGVLTWVGIGAVRSWTKR-----STQQAADSQAVENTLVTA 116  
 QY 125 DTRNRVREVSVDKQVRVWVPMALVLTWL-NENAYLDAFVFIGVGAQY--GDTG-RW 180  
 Db 117 TMAIS-----VGVWPAIRSGIATNLSNPAVL-----FFGSVFAQVREPDGIGW 160  
 QY 181 IFRAGAF--AASLIWFFLVGGAALSRLPS--PKVWRWVWVAVMTALAIAKLMLMG 236  
 Db 161 SIFIGVFLTLTGLWF--VGF--AVLVKLAAGLTNGAIIIDLTGVIFIGLGMFIFEG 216

RESULT 9  
 US-10-374-903A-6  
 ; Sequence 6, Application US/10374903A  
 ; Publication No. US20040038250A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: University of Oviedo  
 ; APPLICANT: Astur Pharma, S.A.  
 ; TITLE OF INVENTION: The gene cluster for thienamycin biosynthesis,  
 ; FILE REFERENCE: Thienamycin-10-AP  
 ; CURRENT APPLICATION NUMBER: US/10/374,903A  
 ; CURRENT FILING DATE: 2003-02-26

; NUMBER OF SEQ ID NOS: 33  
 ; SEQ ID NO 6  
 ; LENGTH: 212  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces cattleya  
 US-10-374-903A-6  
 Query Match 10.8%; Score 128.5; DB 12; Length 212;  
 Best Local Similarity 25.0%; Pred. No. 3.5e-05;  
 Matches 62; Conservative 34; Mismatches 93; Indels 59; Gaps 12;  
 QY 4 MEITITGLL--LGASLLLSI--GPQNVLIKQI--KREGLIAVLLVCLISDVFLF-IAG 56  
 Db 1 MEPMLTATLAFGLGACVLIAAAGPSTMLIIIRQSLHSRRAGFLTVL--GNETGVLTWGVA 58  
 QY 57 TLGVOLLSSNAAPIVLIDIMRWGGIAYLLMFVAVMAAKDMTKVKEAPQIIETEPTVDDTPT 116  
 Db 59 ALGTLTALLAASRTAYDVRIGGAVLVWVYGVQTURAAARG-----EAPSAADD-- 107  
 QY 117 LGSSAVATDTRNRVREVSVDKQVRVWVPMALVLTWLNPNAYLDAFVFIGVGAQYGD 176  
 Db 108 -----EAAVPRSGW-KIYRSGLLLNLANPKAAVAFMSFL----PQFVP 146  
 QY 177 TGR-----WIFAGAPAA-----SLWFFLVGFGAALSRLPSPKVWRWVWVAVV 224  
 Db 147 AGAPKFLPVITAAAPQALFEVGYGYMYVNF-----VGRMKRVISRAGVRRRLQVSGGV 200  
 QY 225 MTALAIAKL 232  
 Db 201 LVLLGIRM 208

RESULT 10  
 US-10-282-122A-77618  
 ; Sequence 77618, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 77618
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77618

Query Match      10.0%; Score 119; DB 12; Length 620;
Best Local Similarity 22.8%; Pred. No. 0.0015;
Matches 63; Conservative 29; Mismatches 72; Indels 112; Gaps 14;

QY 2 VIMEIFITGLLGLSLLSGPQNVIVIKQGIKEGLIA-----VLLVCLISDVFLFI 54
DB 102 MLLGILATGFFGGAIVTP--AMSVLSAVEGI-----AAQDPLAPYVLPFAMMIIVAUFA 155
QY 55 AGTLGVGLLSN-AAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIEETEPTVPD 113
DB 156 VQAMGTERIGRFFAPVM-----LLMFLVLAL----- 181
QY 114 DTPGGSAVADTRNRVRVEVSVDKORVWVKPMMLMAIVLTWLNPNAYLDAPVF----- 166
DB 182 ---LGAHA-----IWHAPQ-----VLRALNP-AYAVHFVLLYQHTL 214
QY 167 -----IGVGAQYDGTGRWIFAGAFASLIWFLPV-----GFGA--AALSR 206
DB 215 FIUGLVLSVGTGVALYADMGHF---GIKPIRIAPFALWPSLLNLYFGQGYLLTISA 270
QY 207 PLSS-----PKVRWNVVAVVMTALAIKMLMG 236
DB 271 PTGSTPFLAPKANLWPLILLATFATVIASQAVISG 306

RESULT 11
US-10-156-761-14994
; Sequence 14994, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14994
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14994

Query Match      9.2%; Score 109.5; DB 14; Length 210;
Best Local Similarity 22.7%; Pred. No. 0.0035;
Matches 53; Conservative 36; Mismatches 107; Indels 37; Gaps 7;

QY 8 ITGLLGASLLSGPQNVIVIKQGIK---REGLIAVLLVCLISDVFLFIAGTLGVGLLS 64
DB 6 VAGLAGYGIAPVGAATVYLSLTARTSLRTGYCAALGVATADGLVALVA-ALGGSALA 64
QY 65 NAAPVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIEETEPTVPDTPGGSAVAT 124
DB 65 AALQPVLPWRWASGLVILALAVRGATAAVERH-YRSPRAAAARRTTTPSP----- 113
```

```
QY 125 DTRNRVRVEVSVDKORVWVKPMMLMAIVLTWLNPNAA--YLDAPVFTGGVGAQ-YGDTGRWI 181
DB 114 -----ARAYLGLLGIITLLNPTVTYFAALV-LGSRAQAQVLPLEQGV 154
QY 182 FAAGAFASLIWFLPVLGFGAALSRPLSSPKVRWNVVAVVMTALAIKML 234
DB 155 FVLAAPFASASQVLLAGGGLGRALTGRGRGLVTALLSSAVILGLAVRLM 207

RESULT 12
US-10-282-122A-44873
; Sequence 44873, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 44873
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-44873

Query Match      8.7%; Score 103.5; DB 12; Length 181;
Best Local Similarity 24.1%; Pred. No. 0.012;
Matches 48; Conservative 35; Mismatches 65; Indels 51; Gaps 10;

QY 42 LLVCLISDVFLFIAGTLGVGLLSNAAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAP 101
DB 19 LQALIQIMIVAA---GVGLFATTPLAFOAVKMFVAYLLYLA----- 60
QY 102 QIIEETPTVPDTPGGSAVATDTRNRVRVEVSVDKORVWVKPMML-AIVLTWLNPNAY 160
DB 61 -YIQWTAP-VKD-----LEIQHEKKOKSALLNGFVNLISNPK-- 98
QY 161 LDAFVFTGGVGAQYDGTGR--WTFAGAFASLIWFLPV---GFG--AALSRPLSGPKV 213
```

Db 99 --ATVELLAVLPQFLDLKSPQIYVL-IMAATVMTVIDLIWAGYTGKASKVLRLRLSPKQ 155  
 QY 214 WRWNVVAVVMTALAIKL 232  
 Db 156 QKYLNGRFVWFSCAALL 174

## RESULT 13

US-10-156-761-10918  
 ; Sequence 10918, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: HATTORI, MASAHIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; PRIOR FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 15109  
 ; SEQ ID NO 10918  
 ; LENGTH: 217  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces avermitilis  
 US-10-156-761-10918

Query Match 8.3%; Score 98.5; DB 14; Length 217;

Best Local Similarity 21.6%; Pred. No. 0.052; 93; Indels 73; Gaps 12;  
 Matches 55; Conservative 34; Mismatches 93;

QY 7 FITGLLGASLLISGPNVLVI-----KQIKREGLLAVLLVCLISDVFLTAGTLGVLD 62  
 Db 10 YLAGLVL---IVLLFGPNLSLVLSVAARKGVRAGYTAAAGVWC--GDTVMTLSAAGVAS 64  
 QY 63 LSNAPVLDIMRWGGIAYLWFAVMAKDA-----MTNKEAPQIIEETPTVPD 113  
 Db 65 LIQANALLFIVKAGAGYLTWLAFGMLRAAMWMTRRORADAAPV----- 113  
 QY 114 DTPLGSSAVATDTRNRVREVSVDKQVWVKPMLMAIVLTWLPNAYLDAFVFIGVGAQ 173  
 Db 114 -APAGE-----RPFRAFVVSLEFNPKAILFFVAFV-----VQ 145  
 QY 174 YQDTRGRWIFAA-----GAFV--ASLIWFLVGFV-----AAALSRLSPSPKVMRWNVV 221  
 Db 146 FVDFPG-YAYFALSFFVLGAFQAQLASFLYLTALIFSGTKLAAAFER---PKRLSAGATTAA 201  
 QY 222 AVVMTALAIKLMJM 236  
 Db 202 GALEFLGFAVKLTLAG 216

## RESULT 14

US-10-282-122A-77782  
 ; Sequence 77782, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 77782  
 ; LENGTH: 295  
 ; TYPE: PRT  
 ; ORGANISM: Versinia pestis  
 US-10-282-122A-77782

Query Match 8.0%; Score 95; DB 12; Length 295;

Best Local Similarity 24.2%; Pred. No. 0.118;  
 Matches 63; Conservative 42; Mismatches 85; Indels 70; Gaps 13;

QY 1 WYMEIFITGLLGASLLISGPNVLVIKQGLK--REGLLAVLLVCLISDVFL-FIGT 57  
 Db 16 ILLIVIAMISIQGASLAKSLFP---LVGAGGITSIRIGITLILFVIFKPRWRFAG- 71  
 QY 58 LGVDLLSNAAPVLDIMRWGGIAYLWFAVMAKDAKMTNKVEAPQIIEETPTVPDTP 117  
 Db 72 -----SRLEPLIYGVALGGMFLFVMSL-----KTVF----- 98  
 QY 118 GGSVAATDTRNRVREVSVDQRV---WYKPMMAIVLTW-LNPAY---LDAFVFIGG 169  
 Db 99 LGIAVALEFTGPLAVAMFSSRAVDFTWVG---LAILGLWFLPLGHTGTIDLFGAACA 155  
 QY 170 VCA-----QYDTRGRWIFAAAGAFASLIWFLP-VGFGAALSRLSPSPKVMR 215  
 Db 156 LGAGACWAIYISGQXAGGDPGTAVGSLTALIFCPIGVANGMALEFPAILP----- 211  
 QY 216 WINVVAVVMTALAIKLM 235  
 Db 212 -IALAVATLSTALPYSLEWV 230

## RESULT 15

US-09-815-242-10183  
 ; Sequence 10183, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.

```
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10183
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10183

Query Match      7.9%; Score 93.5; DB 9; Length 452;
Best Local Similarity 19.4%; Pred. No. 0.47;
Matches 52; Conservative 44; Mismatches 95; Indels 77; Gaps 14;

QY      8 ITGLLG---ASLLSIGPQNVLV-----IKQGIKH--EGLIAVLWCLISDVFLFIAGT 57
Db      150 ISGIILGLVVAEMFTIVRRNVINKLPDVPASVSRSFSLIPGFIILSVNGIILAWLNT 209

QY      58 LG-----VDLLSNAPIVLDIRWGGIAY--LLWFAM---AAKDAMTNKVEAPQIIE 105
Db      210 WGTNPHQIIMDTISTPLASIGSVGCAVVFVFLWFFGIHGALALTALDNGIMTPWALE 269

QY      106 ETEPTVPDDTPLGGSVAT-DTRNRVRVSVVDKQ-RVWVKPMLMAVLVTWLNPNAYLDA 163
Db      270 N-----IATYQYGSVEAALAGKTFHIWAKPM-----LDS 300

QY      164 FVFPGVG-----AQYGDGRWIFAGAFASLLWPELVGFGAALSRL 208
Db      301 FIFLGGGAILGLILAIPTASRRADYQVAKLALFSGIFQ---INEFIL-FGLPIIMNPV 356

QY      209 SSPKYRWINNVVAVVMTALAIKLMNG 236
Db      357 MEIPE-----VLVQPIILAAITLAAAYNG 379
```

Search completed: March 18, 2004, 06:44:48  
Job time : 44 secs

Result No.	Query	Score	Match		Length	DB	ID	Description
			Match	Score				
1	350	29.4	211	1	QHECSA		hypothetical 23K p	
2	347	29.1	211	1	AOH874		probable membrane	
3	343	28.8	211	2	RH1103		hypothetical prote	
4	340	28.5	211	2	R85948		hypothetical prote	
5	335	28.1	205	2	AG0112		probable LysE type	
6	332	27.9	200	2	D83100		probable transport	
7	326.5	27.4	202	2	AG2890		LysE family transp	
8	326.5	27.4	202	2	C97472		probable membrane	
9	319	26.8	201	2	C70744		hypothetical prote	
10	318	26.7	199	2	H07556		hypothetical prote	
11	302.5	25.4	204	2	AD3411		transporter, LysE	
12	245.5	20.6	211	2	D82318		LysE/YggA family p	
13	238	20.0	211	2	S57940		YggA protein hom	
14	176.5	14.8	220	2	D69838		conserved hypothet	
15	173.5	14.6	210	2	F64609		conserved hypothet	
16	170.5	14.3	210	2	F71906		hypothetical prote	
17	164	13.8	205	2	B89854		conserved hypothet	
18	154.5	13.0	205	2	C97219		uncharacterized co	
19	153.5	12.9	213	2	E97789		hypothetical prote	
20	151.5	12.7	208	1	B69066		conserved hypothet	
21	148	12.4	200	2	G83703		hypothetical prote	
22	136.5	11.5	213	2	F83444		hypothetical prote	
23	132	11.1	210	2	B87252		efflux protein, Lys	
24	130	10.9	210	2	D84016		hypothetical prote	
25	120.5	10.1	208	2	G87305		efflux protein, Ly	
26	119	10.0	620	2	F82449		potassium uptake p	
27	117.5	9.9	197	2	B83280		hypothetical prote	
28	115.5	9.7	216	2	AH3203		RhtB family transp	
29	114.5	9.6	208	2	G84086		dihydrodipicolinat	

```

QY      67 APVLIDIMRWGGIAYLLIFANVAAKADAMTKVEAPQLIEETEPTVPDDTPLGSSAVADTT 126
Db      :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
        65 SPMLIALVTWGVVFLTWLWGFAFKTANSSNEL----- 98

QY      127 RNRVRVESVDKQRVVKVMPMLMAIVLTWLPNPNAYLDADFVFIGVGQAQG-DTGRWIFAAG 185
Db      :|::||::||::||::||::||::||::||::||::||::||::||::||:
        99 -----ASAENVLKQGRW-KIIATMLAVTWMHPHYLVTFVVLGSLGGQLDVPKRWFALGT 151

QY      186 AFASLIIFPLVFGGAALSRLSPSPKPVWRINNVAVVMATAIKL 232
Db      :|::||::||::||::||::||::||::||::||::||::||::||::||:
        152 TISASFLEWFGGLAILAALAPLRKTAKSQRIINLVGCVMWFIAQL 198

RESULT 4
E85948
hypothetical protein yggA [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: E85948
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: AB54480; MUID:21074935; PMID:11206551
A/Accession: E85948
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-211 <STO>
A/Cross-references: GB:AE005174; NID:g12517455; PIDN:AAGS8049.1; GSPDB:GN00145; UWGP:Z4
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
C/Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)

Query Match          28.5%; Score 340; DB 2; Length 211;
Best Local Similarity 34.5%; Pred. No. 1-2e-23;
Matches       78; Conservative 37; Mismatches 79; Indels 32; Gaps 2;

QY      7  FITGLLGASLLSIGPQNVIKGIKRGLIANLVCLISDVFLEFIAGTIQVLDLSNA 66
Db      :|::||::||::||::||::||::||::||::||::||::||::||::||:
        5 YFQGLAAGMILPLGFQNFANVQIRRVHMIALLCAISDLVICAGIFGGSALLMQ 64

QY      67 APVLIDIMRWGGIAYLLIFANVAAKADAMTKVEAPQLIEETEPTVPDDTPLGSSAVADTT 126
Db      :|::||::||::||::||::||::||::||::||::||::||::||::||:
        65 SPMLIALVTWGVVFLTWLWGFAFKTANSSNEL----- 98

QY      127 RNRVRVESVDKQRVVKVMPMLMAIVLTWLPNPNAYLDADFVFIGVGQAQGDTGRWIFAAGA 186
Db      :|::||::||::||::||::||::||::||::||::||::||::||::||:
        99 -----ASAENVLKQGRW-KIIATMLAVTWMHPHYLVTFVVLGSLGGQLDVPKRWFALGT 152

QY      187 AFASLIIFPLVFGGAALSRLSPSPKPVWRINNVAVVMATAIKL 232
Db      :|::||::||::||::||::||::||::||::||::||::||::||::||:
        153 ISASFLEWFGGLAILAALAPLRKTAKSQRIINLVGCVMWFIAQL 198

RESULT 5
AG0112
probable LysE type translocator YPO0918 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C/Accession: AG0112
R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.E.;
deno-Farrago, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: AG0112
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-205 <KUR>
A/Cross-references: GB:ALU590842; PIDN:CAC99762.1; FID:g15978989; GSPDB:GN00175
C/Genetics:

```





A:Residues: 1-199 <COL>  
A:Cross-references: GB:Z74025; GB:AL123456; NID:g3261586; PIDN:CA898398.1; PID:gl40346  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: Rv1986  
C:Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)

Query Match 26.7%; Score 318; DB 2; Length 199;  
Best Local Similarity 31.1%; Pred. No. 1.2e-21;  
Matches 70; Conservative 44; Mismatches 77; Indels 34; Gaps 3;

QY 8 ITGLLGASLLSGPQNVLVIKQIGREGLIAVLCISDVFLFTAGTGLGVDLLSNA 67  
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
6 VVGFLACFTLIIAIGAQAFAVLROGIREHVLFPVALCTVSDIVLIAGAIAFGALIGA 65

QY 68 PIVLDIMRWGGIAYLVFWFVAAMKDMTNKVEAPQIIETETPTVPDDTLPGGSATDTR 127  
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
66 FRALNVFKGAAPFLGYGLLAARWRPVALLP-----SGATPVR 106

QY 128 NRRRVEVSVDQRWVKPMLMAIVLTWLNENAYLDAPFIGVGGAQYGTGRWIFAAGAF 187  
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
107 -----LAEVLVTCAAFTPLNFPHVYLDTVLLGALANEHSIQ-RWLFLGAV 151

QY 188 AASLIWFPLVGFGAALSRPLSSPKWRWINVVAVMTALAIKL 232  
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
152 TASAWPFATLGGAGRLGFLTNPGRRIILDGLIAVMVALGISL 196

RESULT 11  
AD3411 transporter, lyse family BMEI1274 [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002  
C:Accession: AD3411  
R:DelVecchio, V.G.; Kaparatil, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.: Mazur, M.; Goldman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:title: The genome sequence of the facultative intracellular pathogen Brucella meliten  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AD3411  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-204 <KUR>  
A:Cross-references: GB:A508917; PIDN:AAL52455.1; PID:gl7983261; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI1274  
A:Map position: I  
A:Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)

Query Match 25.4%; Score 302.5; DB 2; Length 204;  
Best Local Similarity 30.2%; Pred. No. 3.1e-20;  
Matches 68; Conservative 43; Mismatches 81; Indels 33; Gaps 3;

QY 8 ITGLLGASLLSGPQNVLVIKQIGREGLIAVLCISDVFLFTAGTGLGVDLLSNA 67  
DB :|:|:|:|:|:|:|:~::~::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
6 LSGFLGASLIIGAQAFAFLRQGLLRQHVFILCLICALSDALLISAGVAGLGTIAQS 65

QY 68 PIVLDIMRWGGIAYLVFWFVAAMKDMTNKVEAPQIIETETPTVPDDTLPGGSATDTR 127  
DB :|:|:|:|:|:|:|:~::~::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
66 PKLIAEVLTAAGAALFFMYASVAFRAAF--HPEAMQKNSGAVSVSK-----AAAA--- 113

QY 128 NRRRVEVSVDQRWVKPMLMAIVLTWLNENAYLDAPFIGVGGAQYGTGRWIFAAGAF 187  
DB :|:|:|:|:|:|:|:~::~::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
114 -----CLALTFLNPHVYLDTVLLIGLSLARLEGPARAAYGAA 152

QY 188 AASLIWFPLVGFGAALSRPLSSPKWRWINVVAVMTALAIKL 232  
DB :|:|:|:|:|:|:|:~::~::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
153 TASCIFWFGALGVGARLIQPIFAKPAAMRVLDICIGVWMALIGLSL 197

RESULT 12  
B82318

lysE/YggA family protein VC0481 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: B62318  
R:Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
1. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: B62318  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-211 <H21>  
A:Cross-references: GB:AE004134; GB:AE003852; NID:G9654900; PID:AAF93654.1; GSDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC0481  
A:Map position: 1  
C:Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)

Query Match 20.6%; Score 245.5; DB 2; Length 211;  
Best Local Similarity 28.6%; Pred. No. 5e-15; Mismatches 34; Indels 31; Gaps 3;  
Matches 65; Conservative 34

Qy 6 IFITGILLGASLLSIGPQNVLIKQIKREGIAVLVCLISDVFLAGTLGVLLSN 65  
Db 5 ILLQGSIGATMIPIGAQNAVVLNQIKKHHLTTATCGVLDIMFIFLIGFGGALIS 64

Qy 66 AAPVILDMRWGGIAYLLFAVMAAKDAMTKVAPQIETEPTVDDTFLGGSVAATD 125  
Db 65 QNTSLILGVTLAGILFCYGFSLRAAL-----KPPQASESTA-----NPMAGRKAV- 113

Qy 126 TENRVREVSVDQRVWVKPMALVLTWLNPNAYLDAFVFGVGAQYGDGTGWIFAAG 185  
Db 114 -----IFGAPVTVENPHLYLDTVTLGSGGQGGDERISFAIG 153

Qy 186 AFAASLIWFVLGFGAALSRPLSSPKVWKNVWVAVVMTALAIKL 232  
Db 154 TILASFVWFPTLSLGAALKSLTSLKPRVRQVIDMAVAAVMFIAPAL 200

RESULT 13  
S57940  
YggA protein homolog - Aeromonas hydrophila  
C:Species: Aeromonas hydrophila  
C>Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 08-Oct-1999  
C:Accession: S57940  
R:Swift, S.; Fish, L.; Williams, P.; Stewart, G.S.A.B.  
submitted to the EMBL Data Library, July 1995  
A:Reference number: S57938  
A:Accession: S57940  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-211 <SWT>  
A:Cross-references: EMBL:X89469; NID:G899143; PID:CAA61655.1; PID:G899146

Query Match 20.0%; Score 238; DB 2; Length 211;  
Best Local Similarity 26.6%; Pred. No. 2.4e-14; Mismatches 36; Indels 54; Gaps 4;  
Matches 58; Conservative 36

Qy 17 LLSLSIGPQNVLIKQIKREGIAVLVCLISDVFLAGTLGVLLSNAAPVILDMRW 76  
Db 1 MIIPIGAQNAFVLSRGHNRHLLTATLCLDLVIGVFGGANLLAASPGLALLTW 60

Qy 77 GGIAYLLFAVMAAKDAMTK-----VEAPQIETEPTVDDTFLGGSVAATDTRVRV 132  
Db 61 GGVLFLGFGIRSLRGAWRGQAKLADSPQL-----MGYKSV----- 97

Qy 133 EVSVDQRVWVKPMALVLTWLNPNAYLDAFVFGVGAQYGDGTGWIFAAGAFASLI 192  
Db 98 -----LMTGLVTLNPHLYLDTLMLQSGSQPAEELRSFAAVMLASLV 144

Qy 193 WFLPVGFGAALSRPLSSPK-----VWRW 216  
Db 145 WYSLAFAGVAVLSPWLARSQGYSKLILLVSPCWGW 182

RESULT 14  
D69838  
conserved hypothetical protein yisU - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: D69838  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteri  
C.; Bron, S.; Brouillet, S.; Bruchci, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chk  
A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabbret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero  
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, I  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: D69838  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-220 <KUN>  
A:Cross-references: GB:Z99109; GB:AL009126; NID:G2633260; PID:CA812926.1; PID:G2633422  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yisU  
C:Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)

Query Match 14.8%; Score 176.5; DB 2; Length 220;  
Best Local Similarity 22.9%; Pred. No. 1e-08; Mismatches 53; Conservative 36; Indels 43; Gaps 4;  
Matches 53

Qy 4 MEIFTGILLGASLLSIGPQNVLIKQIKREGL---IYVLLVCLISDVFLAGTLGV 60  
Db 18 MNAIHGIVLAFGLILPLGVQNVFFQAGLQKHWRALPAVISAASVCDTLILVAVAGV 77

Qy 61 DLLSNAAPVILDMRWGGIAYLLFAVMAAKDAMTKVAPQIETEPTVDDTFLGGS 120  
Db 78 SVIVQELPFTVMAAGFLFLINGV----- 105

Qy 121 AVATDTRNRVEVSVDQRVWV--KPLMAVLTWLNPNAYLDAFVFGVGAQYGDGTG 178  
Db 106 ----TWN-IRPNTSQNEKHFTTPKQQAFAAAVSLNPHALDITIGVIGTSLQYSGLE 159

Qy 179 RMFPAAGFAASLIWFVLGFGAALSRPLSSPKVWKNVWVAVVMTALA 229  
Db 160 KMLFMAACIAVSWITFISLAIAGRLFQITDITSGRLMLIVNKCSAAVWMAAA 210

RESULT 15  
F64609  
conserved hypothetical integral membrane protein HP0718 - Helicobacter pylori (strain 26  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 17-Mar-2000  
C:Accession: F64609  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney, L.  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: F64609  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

Search completed: March 12, 2004, 18:26:47  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2004, 18:04:50 ; Search time 18 Seconds  
(without alignments)  
682.698 Million cell updates/sec

Title: US-09-105-117K-2

Perfect score: 1191

Sequence: 1 MWMEIFITGLLGASLLS.....INVVAVVMTALAIAKILMLG 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1178	98.9	233	1	LYSE CORGL
2	864.5	72.6	228	1	LYSE COREF
3	350	29.4	211	1	YGAH ECOLI
4	319	26.8	201	1	Y488 MYCTU
5	318	26.7	199	1	Y786 MYCTU
6	310	26.0	206	1	YGAH AERHA
7	254	21.3	225	1	YGAH AERHY
8	105	8.8	206	1	RHTC ECOLI
9	101	8.5	212	1	YEAS ECOLI
10	93.5	7.9	452	1	PTCC ECOLI
11	92.5	7.8	465	1	Y093 RHIME
12	92	7.7	223	1	YAHN ECOLI
13	91	7.6	206	1	RHTC SALTY
14	89	7.5	944	1	CHS2 NEUR
15	88.5	7.4	3137	1	CA36 CHICK
16	88	7.4	206	1	RHTC SALTY
17	87.5	7.3	443	1	DCUA HELPY
18	87.5	7.3	452	1	DNAH STRMU
19	87	7.3	295	1	YBIF SALTY
20	86.5	7.3	195	1	YFIC ECOLI
21	86	7.2	295	1	YBIF ECOLI
22	86	7.2	351	1	MRAY METHA
23	85	7.1	302	1	Y091 METJA
24	84.5	7.1	253	1	CYSZ ECOLI
25	84.5	7.1	443	1	DCUA HELPY
26	84	7.1	482	1	YFIC BACSU
27	84	7.1	685	1	PHUB SALTY
28	83.5	7.0	725	1	HPBA RHOPA
29	83	7.0	216	1	YBIF PSEAE
30	83	7.0	726	1	HPBA THEME
31	82.5	6.9	3491	1	ERV1 SACER
32	81.5	6.8	253	1	CYSZ SALTY
33	81.5	6.8	742	1	DHET ACEAC

Query Match 98.9% Score 1178; DB 1; Length 233;

## ALIGNMENTS

RESULT 1									
ID	LYSE CORGL	STANDARD;	PRT;	233 AA.					
AC	P94633;								
DT	15-JUL-1998 (Rel. 36, Created)								
DT	16-OCT-2001 (Rel. 40, Last sequence update)								
DT	28-FEB-2003 (Rel. 41, Last annotation update)								
DE	Lysine exporter protein.								
GN	LYSE OR CGL1262.								
OS	Corynebacterium glutamicum (Brevibacterium flavum).								
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;								
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.								
OX	NCBI_TaxID=1718;								
RN	[1]								
RP	SEQUENCE FROM N.A., AND FUNCTION.								
RC	STRAIN=RL27;								
RX	MEDLINE=97126810; PubMed=8971704;								
RA	Vrljic M.M., Sahm H., Eggeling L.;								
RT	"A new type of transporter with a new type of cellular function: L-lysine export from Corynebacterium glutamicum.";								
RL	Mol. Microbiol. 22:815-826(1996).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;								
RA	Nakagawa S.;								
RL	*Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";								
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.								
CC	-!- FUNCTION: INVOLVED IN THE EFFLUX OF EXCESS OF L-LYSINE. THIS IS NECESSARY TO CONTROL THE INTRACELLULAR L-LYSINE LEVEL.								
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.								
CC	-!- SIMILARITY: BELONGS TO THE LYSE/YGGA FAMILY.								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).								
CC	EMBL; X96471; CAA65324.2; ..								
CC	EMBL; AP005277; BAB98655.1; ALT INIT.								
DR	InterPro; IPR004777; Lys exporter.								
DR	InterPro; IPR001123; LysE.								
DR	Pfam; PF01810; LysE; 1.								
DR	TIGRFAMs; TIGR00948; 2a75; 1.								
KW	Transport; Transmembrane; Inner membrane; Complete proteome.								
FT	TRANSMEM 3 23 POTENTIAL.								
FT	TRANSMEM 35 55 POTENTIAL.								
FT	TRANSMEM 66 86 POTENTIAL.								
FT	TRANSMEM 144 164 POTENTIAL.								
FT	TRANSMEM 177 197 POTENTIAL.								
FT	TRANSMEM 213 233 POTENTIAL.								
SQ	SEQUENCE 233 AA; 25082 MW; F5FD9B1ACAD1D13 CRC64;								

Q8xw9 escherichia  
P30016 escherichia  
P4504 rhizobium m  
P12610 escherichia  
P12673 salmonella  
Q8f149 escherichia  
Q16572 homo sapien  
P27847 escherichia  
Q8x4v6 escherichia  
P46333 bacillus su  
O35304 mus musculu  
O07708 mycobacteri

34 81 6.8 622 1 KUP ECOLI  
35 81 6.8 622 1 KUP ECOLI  
36 81 6.8 676 1 COMF RHIME  
37 80.5 6.8 253 1 CYSZ ECOLI  
38 80.5 6.8 253 1 CYSZ SALTY  
39 80.5 6.8 456 1 AROP ECOLI  
40 80.5 6.8 532 1 VAT HUMAN  
41 80 6.7 206 1 RHTB ECOLI  
42 80 6.7 394 1 YGAY ECOLI  
43 80 6.7 461 1 CSBC BACSU  
44 80 6.7 530 1 VAT MOUSE  
45 79.5 6.7 215 1 ORN MYCLE

Best Local Similarity 100.0%; Pred. No. 7.1e-95;  
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MEIPIITGLLGLASLLLSIGPQNVLYIKQIKREGIAVLVCLISDVFLFIAGTLGVLL 63  
Db 1 MEIPIITGLLGLASLLLSIGPQNVLYIKQIKREGIAVLVCLISDVFLFIAGTLGVLL 60

QY 64 SNAAPITVLDMRWGIAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVPDTPLGSAVA 123  
Db 61 SNAAPITVLDMRWGIAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVPDTPLGSAVA 120

QY 124 TDTNRNVRVSVKQVWVKPMLMAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVPDTPLGSAVA 183  
Db 121 TDTNRNVRVSVKQVWVKPMLMAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVPDTPLGSAVA 180

QY 184 AGAFAASLIWFLVFGAAALSRPLSSPKVWRNINVVAVMTALAIAKMLMG 236  
Db 181 AGAFAASLIWFLVFGAAALSRPLSSPKVWRNINVVAVMTALAIAKMLMG 233

RESULT 2

LYSE COREF STANDARD; PRT; 228 AA.

AC Q8RQ4;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Lysine exporter protein.  
GN LYSE OR CBL357.  
OS Corynebacterium efficiens.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=152794;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
RA Itaya H., Kimura E., Kawahara Y., Sugimoto S.;  
RT "lysE, lysE of Corynebacterium efficiens";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
RX MEDLINE=22723752; PubMed=12840036;  
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,  
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ieko K.,  
RA Gojohori F.;  
RT "Comparative complete genome sequence analysis of the amino acid  
RT replacements responsible for the thermostability of Corynebacterium  
RT efficiens.";  
RL Genome Res. 13:1572-1579(2003).  
CC -!- FUNCTION: INVOLVED IN THE EFFLUX OF EXCESS OF L-LYSINE. THIS IS  
CC NECESSARY TO CONTROL THE INTRACELLULAR L-LYSINE LEVEL (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE LYSE/YGGA FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AB083133; BAB88827.1; ALT\_INIT.  
CC EMBL; AF005218; BAC16167.1; ALT\_INIT.  
CC InterPro; IPR001123; Lyse.  
CC Pfam; PF01810; Lyse; 1.  
CC Transport; Transmembrane; Inner membrane; Complete proteome.  
DR TRANSMEM 3 23 POTENTIAL.  
DR TRANSMEM 37 57 POTENTIAL.  
DR TRANSMEM 66 86 POTENTIAL.

FT TRANSMEM 139 159 POTENTIAL.  
FT TRANSMEM 172 192 POTENTIAL.  
FT TRANSMEM 208 228 POTENTIAL.  
SQ SEQUENCE 228 AA; 24374 MW; 68FB379A8380EDA7 CRC64;

Query Match 72.6%; Score 864.5; DB 1; Length 228;  
Best Local Similarity 71.2%; Pred. No. 9.1e-68;  
Matches 166; Conservative 30; Mismatches 32; Indels 5; Gaps 2;

QY 4 MEIPIITGLLGLASLLLSIGPQNVLYIKQIKREGIAVLVCLISDVFLFIAGTLGVLL 63  
Db 1 MEIPIITGLLGLASLLLSIGPQNVLYIKQIKREGIAVLVCLISDVFLFIAGTLGVLL 60

QY 64 SNAAPITVLDMRWGIAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVPDTPLGSAVA 123  
Db 61 SNAAPITVLDMRWGIAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVPDTPLGSAVA 115

QY 124 TDTNRNVRVSVKQVWVKPMLMAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVPDTPLGSAVA 183  
Db 116 VTTKQRPRLRITSGTRQVWVRPMLMAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVPDTPLGSAVA 175

QY 184 AGAFAASLIWFLVFGAAALSRPLSSPKVWRNINVVAVMTALAIAKMLMG 236  
Db 176 AGAFAASLIWFLVFGAAALSRPLSSPKVWRNINVVAVMTALAIAKMLMG 228

RESULT 3

YGGA ECOLI STANDARD; PRT; 211 AA.

ID YGGA ECOLI  
AC P11667;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein ygga.  
GN YGGA OR B2923.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474 (1997).  
RN [2]  
RP SEQUENCE OF 15-211 FROM N.A.  
RC STRAIN=K12 / CS520;  
RX MEDLINE=89313302; PubMed=2546007;  
RA Alefounder P.R., Baldwin S.A., Perham S.A., Short N.J.;  
RT "Identification, molecular cloning and sequence analysis of a gene  
RT cluster encoding the class II fructose 1,6-bisphosphate aldolase, 3-  
RT phosphoglycerate kinase and a putative second glyceraldehyde 3-  
RT phosphate dehydrogenase of Escherichia coli.";  
RL Mol. Microbiol. 3:723-732 (1989).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -!- SIMILARITY: BELONGS TO THE LYSE/YGGA FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U28377; AAA69090.1; -.  
CC EMBL; AE000375; AAC75960.1; -.  
CC EMBL; X14436; CAA32607.1; -.

```

DR PIR; B65077; QORCSA.
DR EcoGene; Egil159; YggA.
DR InterPro; IPR004777; Lys exporter.
DR InterPro; IPR001123; LysE.
DR Pfam; PF01810; LysE; 1.
DR TIGRFAMs; TIGR00948; 2a75; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 111 131 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 179 199 POTENTIAL.
SQ SEQUENCE 211 AA; 23175 MW; 2DAFE27B6A9B822 CRC64;

Query Match 29.4%; Score 350; DB 1; Length 211;
Best Local Similarity 35.7%; Pred. No. 2.7e-23;
Matches 81; Conservative 37; Mismatches 75; Indels 34; Gaps 4;

QY 7 FITGLLIGASLLSIGPQNVLVKQIKREGIAVLVCLISDVFLFIAGTLGVDLLSNA 66
DB 5 YFQGLGAAMILPLGPQNAFMQGIHQYHIMALLCAISDLVILICAGIFGGSALIMQ 64
QY 67 APVLIDIMRWGGIAYLLIFAVAAKADMTNKVEAPQIIETEPTVPDTPLGSAVATDT 126
DB 65 SPWLLALVWGGVAFLLWYGGCAKTKAMSNIEL----- 98
QY 127 RNRVREVSVDKQVWVKEMALVLTWLNPNAYLDADFVFGGVGAQYG-DTGRWIFPAAG 185
DB 99 -----ASAEVWKQGRW-KIATMLAVTLNPHVYLDTEVLGSLGGQLDVEPKRW-FALG 151
QY 186 AFASLWFLVFGGAALSRLSPKVVWVWVWVMTALAKL 232
DB 152 TISASFLWFGGLALLAALAPRLTAKAQLINLVGCVWVFALQL 198

RESULT 4
Y488 MYCTU STANDARD; PRT; 201 AA.
AC Q11154;
DT 01-OCT-1996 (Rel. 34, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Hypothetical protein Rv0488/MT0507/Mb0498.
GN Rv0488 OR MT0507 OR MTCY20G9.14 OR Mb0498.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broach R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Bigham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkoeh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

```

```

BA Bishai W., Jacobs W.R. Jr., Venter J.C., Frazer C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempé C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE LYSY/YGGA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z77162; CAB00949.1; -.
CC EMBL; AE006952; AKK44730.1; -.
CC EMBL; BX248335; CAD93361.1; -.
CC PIR; C70744; C70744.
CC TIGR; MT0507; -.
CC TubercuList; Rv0488; -.
CC InterPro; IPR004777; Lys exporter.
CC InterPro; IPR001123; LysE.
CC Pfam; PF01810; LysE; 1.
CC TIGRFAMs; TIGR00948; 2a75; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 25 45 POTENTIAL.
CC FT TRANSMEM 57 77 POTENTIAL.
CC FT TRANSMEM 104 124 POTENTIAL.
CC FT TRANSMEM 133 153 POTENTIAL.
CC FT TRANSMEM 169 189 POTENTIAL.
CC SQ SEQUENCE 201 AA; 20951 MW; E198975DF088B6E4 CRC64;

Query Match 26.8%; Score 319; DB 1; Length 201;
Best Local Similarity 32.4%; Pred. No. 1.2e-20;
Matches 71; Conservative 50; Mismatches 64; Indels 34; Gaps 4;

QY 16 SLLSIGPQNVLVKQIKREGIAVLVCLISDVFLFIAGTLGVDLLSNAPIVLDIMR 75
DB 3 TLKVAIGPQNAVLRQIGIRREYVLVIVALGCIADGALIAAGVGGFAALIIHAHPNMTLVAR 62
QY 76 WGGIAYLLWFAVMAKADMTNKVEAPQIIETEPTVPDTPLGSAVATDTRNRVREVS 135
DB 63 FGGAFLIGYALLAARNW-----RPSGLVPSES--GPAAL----- 96
QY 136 VDKQRVWVKPMALVLTWLNPNAYLDADFVFGGVGAQYGDTRWIFPAAGFAASLIWFP 195
DB 97 -----IGTVQMLCVTFPLNPHVYLDTEVLIGLANEESDL-RWFFGAGAAVWVFA 148
QY 196 LVGFGAALSRLSPKVVWVWVWVMTALAKML 234
DB 149 VLFGSAGRLQPPFATPAWRILDLAVMTVMIGVAVVVLV 187

RESULT 5
YJ86 MYCTU STANDARD; PRT; 199 AA.
AC Q10871;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein Rv1986/MT2040/Mb2008.
GN Rv1986 OR MT2040 OR MTCY39.33C OR Mb2008.

```

OS Mycobacterium tuberculosis, and  
 OS Mycobacterium bovis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773, 1765;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.tuberculosis; STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,  
 RA Sultana J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence";  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
 RX MEDLINE=22206494; PubMed=12218036;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayan L.A., Ermolaeva M., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains";  
 RL J. Bacteriol. 184:5479-5490(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.bovis; STRAIN=AF2122/97;  
 RX MEDLINE=22709107; PubMed=12788972;  
 RA Garnier T., Eigmler K., Camus J.-C., Medina N., Mansoor H.,  
 RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
 RA "The complete genome sequence of Mycobacterium bovis";  
 RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).  
 RL [4]  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE LYSE/YGGA FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; E74025; CAA98398.1; -;  
 CC EMBL; AE007056; AK46315.1; -;  
 CC EMBL; BX248341; CAD94861.1; -;  
 CC PIR; H70756; H70756.  
 CC TIGR; MT2040; -;  
 CC Tuberculin; RV1986; -;  
 CC InterPro; IPR004777; Lys exporter.  
 CC InterPro; IPR001123; Lyse.  
 CC Pfam; PF01810; Lyse; 1.  
 CC TIGRfam; TIGR00948; 2a75; 1.  
 CC Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 6 26 POTENTIAL.  
 FT TRANSMEM 42 62 POTENTIAL.  
 FT TRANSMEM 68 88 POTENTIAL.  
 FT TRANSMEM 116 136 POTENTIAL.  
 FT TRANSMEM 144 164 POTENTIAL.  
 FT TRANSMEM 178 198 POTENTIAL.  
 SQ SEQUENCE 199 AA; 20775 MW; 6F330132DCFD0FF CRC64;  
 Query Match 26.7%; Score 318; DB 1; Length 199;  
 Best Local Similarity 31.9%; Pred. No. 7.6e-20;  
 Matches 72; Conservative 43; Mismatches 77; Indels 34; Gaps 3;  
 QY 8 ITGLLGASLLLSIGPQNVIVIKQIKREGILAVLVCLISDVFLFIAGTGVLDLSNAA 67

Best Local Similarity 31.1%; Pred. No. 1.5e-20;  
 Matches 70; Conservative 44; Mismatches 77; Indels 34; Gaps 3;  
 QY 8 ITGLLGASLLLSIGPQNVIVIKQIKREGILAVLVCLISDVFLFIAGTGVLDLSNAA 67  
 Db 6 VVGFLACFTLIAAIGAGQNAFVLRQGIQREHVPVVALCTVSDIVLIAAGIAGFALIGAH 65  
 QY 68 PIVLDIMRWAGGIAYLJAFVAAAKDAMTKVEAPQIIEETEPTVDDTPLGGSAAVADTR 127  
 Db 66 PRALNVVKKGGAAFLTIGYGLAARARVVALIP-----SGATPVR 106  
 QY 128 NRVRVEVSDKQVWVWVPMALVFLVNLNAYLDAFVFGVGAGYGTGRWTFAGAF 187  
 Db 107 -----LAEVLVTCAAFTFLNPHVYLDTVVLGALANEHSDQ-RWLFGLGAV 151  
 QY 188 AASLIWFLVGGGAALSRPLSSPKVWVWVAVVMTALA1KL 232  
 Db 152 TASAVWFATLFGAGRLGFLTNPGSWRILDLGLIAVMVALGISL 196  
 RESULT 6  
 YGGA AERSA STANDARD; PRT; 206 AA.  
 ID YGGA AERSA AC P70775;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein YGGA.  
 GN YGGA.  
 OS Aeromonas salmonicida.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;  
 OC Aeromonadaceae; Aeromonas.  
 OX NCBI\_TaxID=645;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCIMB 1102;  
 RX MEDLINE=97431471; PubMed=9286976;  
 RA Swift S., Karlyshev A.V., Fish L., Durant E.L., Winslow M.K.,  
 RA Chhabra S.R., Williams P., Macintyre S., Stewart G.S.A.B.;  
 RA "Quorum sensing in Aeromonas hydrophila and Aeromonas salmonicida:  
 RT identification of the LuxRI homologs AhyRI and AaARI and their  
 RT cognate N-acylhomoserine lactone signal molecules";  
 RL J. Bacteriol. 179:5271-5281(1997).  
 RL [2]  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE LYSE/YGGA FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U65741; AAB70019.1; ALT\_INIT.  
 CC InterPro; IPR004777; Lys exporter.  
 CC InterPro; IPR001123; Lyse.  
 CC Pfam; PF01810; Lyse; 1.  
 CC TIGRfam; TIGR00948; 2a75; 1.  
 CC Hypothetical protein; Transmembrane.  
 FT TRANSMEM 1 21 POTENTIAL.  
 FT TRANSMEM 37 57 POTENTIAL.  
 FT TRANSMEM 65 85 POTENTIAL.  
 FT TRANSMEM 116 136 POTENTIAL.  
 FT TRANSMEM 148 168 POTENTIAL.  
 FT TRANSMEM 185 205 POTENTIAL.  
 SQ SEQUENCE 206 AA; 21505 MW; DIC2C492CDA0179A CRC64;  
 Query Match 26.0%; Score 310; DB 1; Length 206;  
 Best Local Similarity 31.9%; Pred. No. 7.6e-20;  
 Matches 72; Conservative 43; Mismatches 77; Indels 34; Gaps 3;  
 QY 8 ITGLLGASLLLSIGPQNVIVIKQIKREGILAVLVCLISDVFLFIAGTGVLDLSNAA 67

```

Db      6  LQFTGLGAMIPIGAQNAFVLSRGIRHNHLLAATLCCCLILIGVFGGANLLAAS 65
QY      68  PIVLDMRWGGIAYLLWFVAAMKADMTNKVEAPQIIEETPTVPDDTPIGGSAVATDTR 127
Db      66  PIGLALLTWGGVFLCNGFIRSLSAQGGQ-----GAALADGFR 104
QY      128  NRVREVSVDKQVRWKPKML-VAIVLTLNPNAYLDAFVIGGCAQYGTGTGWIFPAAGA 186
Db      105  -----LMGVKSVLMTLGVTLNPHVLTMLLGSFGSGFAPLRPAAGA 152
QY      187  FAASLWFLPFGGAALSRPLSPKWRWNVVAVMTALAIAKL 232
Db      153  MLASLWVFLAFGAALSPWLARGVQQAIDTIVGLMGLALQL 198

RESULT 7
YCGA AERHY STANDARD; PRT; 225 AA.
AC P52047;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Hypothetical 24.5 kDa protein in ahyr-cdpD intergenic region.
OS Aeromonas hydrophila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;
OC Aeromonadaceae; Aeromonas.
OX NCBI_TaxID=644;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Al;
RX MEDLINE=97431471; PubMed=9286976;
RA Swift S., Kariyeh A.V., Fish L., Durant E.L., Winson M.K.,
RA Chhabra S.R., Williams P., Macintyre S., Stewart G.S.A.B.;
RT "Quorum sensing in Aeromonas hydrophila and Aeromonas salmonicida:
RT identification of the LuxI homologs AhylR and AhylI and their
RT cognate N-acylhomoserine lactone signal molecules";
RL J. Bacteriol. 179:5271-5281(1997).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -I- SIMILARITY: BELONGS TO THE LYSE/YCGA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X89469; CAA61655.1; ALT INIT.
DR InterPro; IPR004777; Lys exporter.
DR InterPro; IPR001123; LysE.
DR Pfam; PF01810; LysE; 1.
DR TIGRFAMs; TIGR00948; 2a75; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 116 136 POTENTIAL.
FT TRANSMEM 150 170 POTENTIAL.
SQ SEQUENCE 225 AA; 24482 MW; 172DB1044739B089 CRC64;

Query Match
Best Local Similarity 26.98; Pred. No. 5.8e-15;
Matches 61; Conservative 38; Mismatches 74; Indels 54; Gaps 4;

QY      8  ITGLLAGSLISGPQNLVLIKQIGREGIAYLVCLISDVFLFIAGTLGVLLSNAA 67
Db      6  LQFTGLGAMIPIGAQNAFVLSRGIRHNHLLAATLCCCLDLVIGVFGGANLLAAS 65
QY      68  PIVLDMRWGGIAYLLWFVAAMKADMTNK-----VEAPQIIEETPTVPDDTPIGGSAVA 123
Db      66  PIGLALLTWGGVFLCNGFIRSLSRAWGGAKLADSPQ-----MGVKS- 111

```

```

QY      124  TDTNRVRVSVSKQVRWKPKMLVAIVLTLNPNAYLDAFVIGGCAQYGTGTGWIFA 183
Db      112  -----LMTLGVTLNPHVLTMLLGSFGSGFAELASAF 149
QY      184  AGAFAASLWFLPFGGAALSRPLSSPK-----VWR 216
Db      150  AVAMLASLWVFLAFGAALSPWLARGVQQAIDTIVGLMGLALQL 196

RESULT 8
RHTC ECOLI STANDARD; PRT; 206 AA.
AC P27846;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Threonine efflux protein.
GN RHTC OR B3823 OR Z5344 OR ECS4753.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=92358234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
RT from 84.5 to 86.5 minutes.";
RL Science 257:771-778(1992).
RN [2]
RP REVISIONS.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [5]
RP SEQUENCE OF 1-107 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=87115164; PubMed=3027506;
RA Irino N., Nakayama K., Nakayama H.;
RT "The reqQ gene of Escherichia coli K12: primary structure and
RT evidence for SOS regulation.";
RL Mol. Gen. Genet. 205:298-304(1986).

```



## RESULT 9

QY 179 RWIFAA-----GAPAA-----SLWFLVGFCA 201  
 DB 153 PFILATLELVSCYLSFLIISGAFVQVIRTKKLAKVGNLSLGLMFVGFCA 205

RESULT 10  
 PTCC\_ECOLI STANDARD; PRT; 452 AA.  
 AC P17334; P76212; P76907; P77332;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE PHS system, N,N'-diacetylchitobiose-specific IIC component (EIIIC-Chb)  
 DE (N,N'-diacetylchitobiose-permease IIC component) (Phosphotransferase  
 DE enzyme II, C component).  
 GN CHBC OR CELB OR B1737.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X12;  
 RX MEDLINE=90185127; PubMed=2179047;  
 RA Parker L.L., Hall B.G.;  
 RT "Characterization and nucleotide sequence of the cryptic cel operon  
 RT of Escherichia coli K12.";  
 RL Genetics 124:455-471 (1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474 (1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X12;  
 RX MEDLINE=97251357; PubMed=9097039;  
 RA Alba T., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,  
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,  
 RA Sempel G., Seki Y., Sivasaundaram S., Tagami H., Takeda J.,  
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;  
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 28.0-40.1 min region on the linkage map.";  
 RL DNA Res. 3:363-377 (1996).  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=91227627; PubMed=2092358;  
 RA Reizer J., Reizer A., Sailer M.H. Jr.;  
 RT "The cellobiose permease of Escherichia coli consists of three  
 RT proteins and is homologous to the lactose permease of Staphylococcus  
 RT aureus.";  
 RL Res. Microbiol. 141:1061-1067 (1990).  
 RN [5]  
 RP IDENTIFICATION OF CHB OPERON.  
 RX MEDLINE=98070757; PubMed=9405618;  
 RA Keyhani N.O., Roseman S.;  
 RT "Wild-type Escherichia coli grows on the chitin disaccharide,  
 RT N,N'-diacetylchitobiose, by expressing the cel operon.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:14367-14371 (1997).  
 RN [6]  
 RP FUNCTION.  
 RX MEDLINE=20490752; PubMed=10913117;  
 RA Keyhani N.O., Wang L.-X., Lee Y.C., Roseman S.;  
 RT "The chitin disaccharide, N,N'-diacetylchitobiose, is catabolized by  
 RT Escherichia coli and is transported/phosphorylated by the  
 RT phosphoenolpyruvate:glycose phosphotransferase system.";

J. Biol. Chem. 275:33084-33090 (2000).  
 CC -!- FUNCTION: This is a component of the phosphoenolpyruvate-dependent  
 CC sugar phosphotransferase system (PTS), a major carbohydrate active  
 CC -transport system. The IIC domains contain the sugar binding site  
 CC and the transmembrane channel; the IIA domain contains the primary  
 CC phosphorylation site (the donor is phospho-HPr); IIA transfers its  
 CC phosphoryl group to the IIB domain which finally transfers it to  
 CC the sugar.  
 CC -!- PATHWAY: N,N'-diacetylchitobiose utilization.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
 CC -!- INDUCTION: By N,N'-diacetylchitobiose.  
 CC -!- SIMILARITY: Contains 1 PTS EIIIC domain.  
 CC -!- CAUTION: Was originally (Ref.4) characterized as part of a cryptic  
 CC cel operon for a cellobiose degradation system. The Cel<sub>1</sub> phenotype  
 CC is due to mutations making expression chitobiose-independent and  
 CC altering the substrate specificity.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; X52890; CAA37070.1; -;  
 CC EMBL; AE000268; AAC74807.1; -;  
 CC EMBL; D90816; BAA15518.1; -;  
 CC EMBL; D90817; BAA15526.1; -;  
 CC EMBL; D90818; BAA15532.1; -;  
 CC PIR; A64933; A64933.  
 CC PIR; S10871; S10871.  
 CC EcoGene; EG10141; chbC.  
 CC InterPro; IPR004796; Cello\_Pts\_IIC.  
 CC InterPro; IPR003352; Ptrans\_EIIC.  
 CC InterPro; IPR004501; Pts\_lac.  
 CC Pfam; PF02378; Pts\_EIIC; 1.  
 CC TIGRFAMs; TIGR00359; Cello\_Pts\_IIC; 1.  
 CC TIGRFAMs; TIGR00410; lacB; 1.  
 CC Phosphotransferase system; Sugar transport; Transmembrane;  
 CC Inner membrane; Complete proteome.  
 CC TRANSMEM 35 55 POTENTIAL.  
 CC TRANSMEM 86 106 POTENTIAL.  
 CC TRANSMEM 114 134 POTENTIAL.  
 CC TRANSMEM 147 167 POTENTIAL.  
 CC TRANSMEM 187 207 POTENTIAL.  
 CC TRANSMEM 229 249 POTENTIAL.  
 CC TRANSMEM 300 320 POTENTIAL.  
 CC TRANSMEM 344 364 POTENTIAL.  
 CC TRANSMEM 367 387 POTENTIAL.  
 CC TRANSMEM 409 429 POTENTIAL.  
 CC CONFLICT 190 205 LIPGFIILSVNGIAT -> FNSRLVYSFRDGYCL (IN  
 CC REF. 1).  
 CC CONFLICT 233 235 VGM -> WL (IN REF. 1).  
 CC CONFLICT 240 265 FVLLVFFGHGALATLALDNGIMTP -> LSTALVIRIHA  
 CC ACADRTGQHYDA (IN REF. 1).  
 CC CONFLICT 361 366 FVLVQP -> LYVYNR (IN REF. 1).  
 CC CONFLICT 405 452 SVALIIVALENLGIATLILPVVYANKAQNADKESSED  
 CC IANALKF -> TSPHCHWSHSTLASQR (IN REF. 1).  
 CC SEQUENCE 452 AA; 48332 MW; EC3CBFEED6231068 CRC64;  
 Query Match 7.9%; Score 93.5; DB 1; Length 452;  
 Best local Similarity 19.4%; Pred. No. 0.9;  
 Matches 52; Conservative 44; Mismatches 95; Indels 77; Gaps 14;  
 QY 8 ITGLLIG--ASLLLSIGQNVLV-----IKQIKR--EGLIIVLVCLISDYFLIAGT 57  
 DB 150 ISGIIGLVWAEFTFIVRNWVIXLPDPSVPSVFSALIPGFIILSVNGIATL 209  
 QY 58 LG-----VLLSNAAPIVLDMRWGGIAY--LLWFAVW--AAKDAMTKVEAFQIE 105  
 DB 210 WGTNFHQIIMDTITPLASIGSVGWAVIVFVLLWFFGHGALATLALONGIWTWALE 269

QY 106 ETEPTVDDTGLGSAVAT-DTNRVRVEVSVDKQ-RVMVKPMLMAIVLTLNPNAYLDA 163  
 Db 270 N-----IATYQGVSEALAAAGTKTHWAKM-----LDS 300  
 QY 164 FVFIGGVG-----AQYDGTGRWIFAFAGAFASLIWFFLVFGGAAALSRPL 208  
 Db 301 FIFLGGSGATLGLILAFIASFARADYRVQAKLPSGIFQ---INEPIL-FGLPIIMNFV 356  
 QY 209 SSFQVRWVNVVAVMTALAKMLMG 236  
 Db 357 MFIPP-----VLVQPILAAITLAAYTMG 379

RESULT 11  
 Y093\_RHIME  
 ID Y093\_RHIME STANDARD; PRT; 465 AA.  
 AC 087394;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical transport protein R00093.  
 GN R00093 OR SMC02616.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RA Powers E.L.; Vuyyuru V., Kahn M.L.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capella D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ransperger U.,  
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 Sinorhizobium meliloti strain 1021."  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 RL -1- FUNCTION: Probable amino-acid or metabolite transport protein.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: Belongs to the amino acid permease family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF055582; AAC62224.1; -  
 CC EMBL; AL591782; CAC41480.1; -  
 CC InterPro; IPR002293; AA/rel permease1.  
 CC InterPro; IPR004840; AAC permease.  
 CC InterPro; IPR004841; Permease region.  
 CC Pfam; PF00324; aa\_permeases; 1.  
 CC PROSITE; PS00218; AMINO ACID PERMEASE 1; FALSE NEG.  
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.  
 FT TRANSNMEN 19 POTENTIAL.  
 FT TRANSNMEN 50 70  
 FT TRANSNMEN 91 111  
 FT TRANSNMEN 140 160  
 FT TRANSNMEN 164 184  
 FT TRANSNMEN 201 221  
 FT TRANSNMEN 244 264  
 FT TRANSNMEN 288 308  
 FT TRANSNMEN 342 362  
 FT TRANSNMEN 363 383

FT TRANSNMEN 403 423 POTENTIAL.  
 SQ SEQUENCE 465 AA; 50783 MW; 0375E164F737AA0A CRC64;  
 Query Match 7.8%; Score 92.5; DB 1; Length 465;  
 Best Local Similarity 20.0%; Pred. NO. 1.1;  
 Matches 47; Conservative 38; Mismatches 83; Indels 67; Gaps 7;  
 QY 17 LLLSIGPQNVLVKQGI-----KREGLIIVLVCLISDFVLFAGTFLGVDLL 63  
 Db 16 LLRVLGFHFWALGVIVLVEYMGWVSVGKGMVAGLMACWVAGLLYTCVAMIDSEVT 75  
 QY 64 SNAAP-----IVLDIMRWG-GIAYLLWFVMAAKDMKNKVEAPQIIEETETVP 112  
 Db 76 STVAAGGQYQAQKHIVGLPMFNVGLFLVMAYTLEANAAT----- 118  
 QY 113 DDTPLGSAVATDTNRVRVEVSVDKQ-RVMVKPMLMAIVLTLNPNAYLDAFVFIGVGVA 172  
 Db 119 -----VGFLDVTAGVGGQGTGLNQPPV---LAIWFLAMLVNTRGVLATYF----- 162  
 QY 173 QYDGTGRWIFAFAGAFASLIWFFLVFGGAAAL-----SRPLSSPKVWRWNVVAV 223  
 Db 163 -----NLVITAIATLAIVLFSVQFGASAVPLDFAITSDPLPYGWWGVASLI 211

RESULT 12  
 YAHN\_ECOLI  
 ID YAHN\_ECOLI STANDARD; PRT; 223 AA.  
 AC P75693; P71307;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein yahn.  
 GN YAHN OR B0328.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12."  
 RL Science 277:1453-1474(1997).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,  
 RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H.,  
 RA Lin D., Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: Belongs to the rht family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE000140; AAC73431.1; -  
 CC EMBL; U73857; AAB18053.1; -  
 CC PIR; H64759; H64759.  
 CC Ecogene; EGI3558; yahn.  
 CC InterPro; IPR004778; Homoser\_Thr\_eff.  
 CC InterPro; IPR001123; Lyse.  
 CC Pfam; PF01810; Lyse; 1.  
 CC TIGRFAMs; TIGR00949; 2A76; 1.  
 KW Hypothetical protein; Transmembrane; Complete proteome.



```

CC -----
DR EMBL; X77782; CAA54816.1; -
DR EMBL; M82951; AAA33592.1; -
DR PIR; B45189; T47246.
DR PIR; T47246; T47246.
DR InterPro; IPR004834; Chitin synth.
DR Pfam; PF01644; Chitin synth; 1.
DR ProDom; PD002998; Chitin synth; 1.
KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
KW Multigene family.
FT TRANSMEM 597 617 POTENTIAL.
FT TRANSMEM 634 654 POTENTIAL.
FT TRANSMEM 669 689 POTENTIAL.
FT TRANSMEM 713 733 POTENTIAL.
FT TRANSMEM 873 893 POTENTIAL.
FT SEQUENCE 944 AA; 106816 MW; F70052AE0803060D CRC64;
SQ
Query Match 7.5%; Score 89; DB 1; Length 944;
Best Local Similarity 20.3%; Pred. No. 4.7;
Matches 60; Conservative 54; Mismatches 95; Indels 86; Gaps 14;
Qy 7 FITGILLGASLLISGPONVIVIKQIGKEGLIAVLVCLT-----SDVLFITAG----- 56
Db 641 YVCVLLICTQFISLGNR-----POGAKMYLASMIIYAVIMVYTTFTFIVVRQIQPS 695
Qy 57 -----TLGVLLSNAAPIVLIDIMRWGGIAYLLMFVAAKADAMTKVE-----APQI 103
Db 696 QKSDDKPDLNGLNVFTN---LIVSVASTLGLYFVMSFLYLDPMWMTSAIQYFVLPSY 752
Qy 104 I-----ETETVDD---TPIGGSVAATDT----- 126
Db 753 ICTLIQIYAFCTHDVTWCTKGDVNMRTDLGGAIVKGSVELEMPDQLDIPDSGYDECLRN 812
Qy 127 -RNRVY---EYSDVK-ORVWVKDMLMAVLWLNPNAYLDADFVFGVQYQYGDTRWI 181
Db 813 LRDKWVPCKSUSEQDQDYKSVRTYVWVSWVATLAM-----AVSAYGDSF--- 864
Qy 182 FAAGAFASLIIFPLVFGGAALSRPLSSPKVWRWVNVVAVVMTALAKMLMG 236
Db 865 IGDNYLRFLTM-----AVALALFRLGSS-TTFAALNVSLVAGVRVRLNMGK 914
RESULT 15
ID _CA36 CHICK STANDARD; PRT; 3137 AA.
AC P15959;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Collagen alpha 3(VI) chain precursor.
GN COL6A3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE OF 1-853 FROM N.A.
RC TISSUE=orta;
EX MEDLINE=91035630; PubMed=1977751;
RA Doliana R., Bonaldi P., Colombatti A.;
RT "Multiple forms of chicken alpha 3(VI) collagen chain generated by
RT alternative splicing in type A repeated domains.";
RN J. Cell Biol. 111:2197-2205(1990).
RP [2]
RP SEQUENCE OF 224-2871 FROM N.A.
RX MEDLINE=90212613; PubMed=2322559;
RA Bonaldi P., Russo V., Buccioti F., Doliana R., Colombatti A.;
RT "Structural and functional features of the alpha 3 chain indicate a
RT bridging role for chicken collagen VI in connective tissues.";
RL Biochemistry 29:1245-1254(1990).
RN [3]
RP SEQUENCE OF 2871-3137 FROM N.A.

```

```

RX MEDLINE=90062147; PubMed=2584214;
RA Bonaldi P., Colombatti A.;
RT "The carboxyl terminus of the chicken alpha 3 chain of collagen VI is
RT a unique mosaic structure with glycoprotein Ib-like, fibronectin type
RT III, and Kunitz modules.";
RL J. Biol. Chem. 264:20235-20239(1989).
CC -I- FUNCTION: Collagen VI acts as a cell-binding protein.
CC -I- SUBUNIT: Trimers composed of three different chains: alpha 1(VI),
CC alpha 2(VI), and alpha 3(VI).
CC -I- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=1;
CC Comment=At least 2 isoforms are produced;
CC Name=1;
CC IsoId=P15989-1; Sequence=Displayed;
CC -I- PTM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -I- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC -I- SIMILARITY: Contains 1 fibronectin type III domain.
CC -I- SIMILARITY: Contains 12 VWFA domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M24282; AAA03201.1; -.
DR PIR; A37797; A37797.
DR HSP; P12111; 2XMT.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR002223; Kunitz_BPTI.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01391; Collagen; 6.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR Pfam; PF00092; vwa; 11.
DR PRINTS; PR00759; BASICPTASE.
DR PRINTS; PR00453; VWFADOMAIN.
DR ProDom; PD000007; Clg_helix; 1.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00131; KU; 1.
DR SMART; SM00327; VWA; 12.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.
DR PROSITE; PS50234; VWFA; 12.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Cell adhesion; Serine protease inhibitor; Collagen;
KW Signal; Alternative splicing.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 3137 COLLAGEN ALPHA 3(VI) CHAIN.
FT DOMAIN 26 2042 NONHELICAL REGION.
FT DOMAIN 2043 2379 TRIPLE-HELICAL REGION.
FT DOMAIN 2380 3137 NONHELICAL REGION.
FT DOMAIN 38 212 VWFA 1.
FT DOMAIN 241 418 VWFA 2.
FT DOMAIN 444 623 VWFA 3.
FT DOMAIN 644 817 VWFA 4.
FT DOMAIN 842 1014 VWFA 5.
FT DOMAIN 1035 1207 VWFA 6.
FT DOMAIN 1239 1410 VWFA 7.
FT DOMAIN 1441 1621 VWFA 8.
FT DOMAIN 1641 1814 VWFA 9.
FT DOMAIN 1840 2029 VWFA 10.
FT DOMAIN 2407 2587 VWFA 11.
FT DOMAIN 2625 2821 VWFA 12.
FT DOMAIN 2945 3043 FIBRONECTIN TYPE-III.
FT DOMAIN 3068 3137 BPTI/KUNITZ INHIBITOR.
FT SITE 2166 2172 INTERRUPTION IN COLLAGENOUS REGION.

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2004, 18:20:26 ; Search time 45 Seconds  
(without alignments)  
1654.718 Million cell updates/sec

Title: US-09-105-117K-2  
Perfect score: 1191  
Sequence: 1 VYMEIFITGLIGALLS.....INVVVVVTAIAIKMLMG 236

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacterioph.\*  
17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	29.3	202	16 Q8XT9	Q8xt9 ralstonia s
2	347	29.1	211	16 Q8FE55	Q8fe55 escherichia
3	347	29.1	211	16 Q8Z3W2	Q8z3w2 salmonella
4	346	29.1	200	16 Q87X28	Q87x28 pseudomonas
5	346	29.1	211	16 Q8ZM68	Q8zm68 salmonella
6	343	28.8	211	16 Q8XD10	Q8xd10 escherichia
7	343	28.8	211	16 Q7UBP8	Q7ubp8 shigella fl
8	336	28.2	212	16 Q8Y2F1	Q8y2f1 ralstonia s
9	335	28.1	205	16 Q8ZHH6	Q8zhb6 yersinia pe
10	332	27.9	200	16 Q8HW36	Q8hw36 pseudomonas
11	328.5	27.6	204	16 Q9K4K6	Q9k4k6 streptomyces
12	328	27.5	208	16 Q8PHW5	Q8phw5 xanthomonas
13	326.5	27.4	202	16 Q8UGV8	Q8ugv8 agrobacteri
14	323	27.1	204	16 Q8APD4	Q8apd4 pseudomonas
15	309	25.9	208	16 Q8P6L3	Q8p6l3 xanthomonas
16	302.5	25.4	204	16 Q8YGB7	Q8ygb7 brucella me

17	301	25.3	203	16 Q82P13	Q82p13 streptomyces
18	296	24.9	206	16 Q8ED97	Q8ed97 shewanella
19	291.5	24.5	204	16 Q8G1N8	Q8g1n8 brucella su
20	285	23.9	226	16 Q9CKJ7	Q9ckj7 pasteurella
21	282.5	23.7	209	16 Q87LL5	Q87ll5 vibrio para
22	281	23.6	186	16 Q83J77	Q83jt7 shigella fl
23	275.5	23.1	206	16 Q8DC97	Q8dc97 vibrio vuln
24	266	22.3	211	16 Q7VS78	Q7vs78 bordetella
25	263	22.1	211	16 Q7WQU0	Q7wqu0 bordetella
26	263	22.1	211	16 Q7W1W4	Q7w1w4 bordetella
27	245.5	20.6	211	16 Q9KUN4	Q9kun4 vibrio chol
28	176.5	14.8	220	16 Q86730	Q86730 bacillus su
29	175.5	14.7	207	16 Q8RHK9	Q8rnx9 fusbacteri
30	173.5	14.6	210	16 Q25420	Q25420 helicobacte
31	170.5	14.3	210	16 Q9ZLC4	Q9zlc4 helicobacte
32	164	13.8	205	16 Q99V11	Q99v11 staphylococ
33	162.5	13.6	205	16 Q81E18	Q81e18 bacillus ce
34	160	13.4	205	16 Q8NX12	Q8nx12 staphylococ
35	158.5	13.3	204	16 Q81RQ2	Q81rq2 bacillus an
36	158	13.3	193	16 Q88HC2	Q88hc2 pseudomonas
37	157.5	13.2	202	16 Q88X38	Q88x38 lactobacill
38	154.5	13.0	205	16 Q97FY1	Q97fy1 clostridium
39	153.5	12.9	213	16 Q92H04	Q92hq4 rickettsia
40	151.5	12.7	208	17 Q97538	Q97538 methanobact
41	148	12.4	200	16 Q9KFP7	Q9kfp7 bacillus ha
42	143	12.0	207	16 Q8LAX1	Q8lax1 bacillus ce
43	136.5	11.5	197	16 Q8EG90	Q8eg90 shewanella
44	136.5	11.5	206	16 Q8Y2B8	Q8y2b8 ralstonia s
45	136.5	11.5	213	16 Q9I3A2	Q9i3a2 pseudomonas

## ALIGNMENTS

RESULT 1  
Q8XT9 PRELIMINARY; PRT; 202 AA.  
ID Q8XT9  
AC Q8XT9; 01-MAR-2002 (TREMELrel. 20, Created)  
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Probable transmembrane protein.  
GN RSC2024 OR RS03592.  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Ralstonia.  
OX NCBI\_TaxID=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GMI1000;  
RX MEDLINE=21681879; PubMed=11823852;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,  
Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,  
Weissenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
RL Nature 415:497-502(2002).  
DR EMBL; AL646067; CADI5726.1; -;  
DR GO; GO:0016020; C-membrane; IEA.  
DR GO; GO:0005293; F1yline permease activity; IEA.  
DR GO; GO:0006865; P-amino acid transport; IEA.  
DR InterPro; IPR001123; LysE.  
DR Pfam; PF01810; LysE; 1.  
KW Complete proteome.  
SQ SEQUENCE 202 AA; 21850 MW; 74346064388E353 CRC64;

Query Match 29.3%; Score 349; DB 16; Length 202;  
Best Local Similarity 35.1%; Pred. No. 1.6e-22;  
Matches 80; Conservative 40; Mismatches 72; Indels 36; Gaps 5;  
QY 10 GLLIGALLSITGPNVIVIKQIKREGIIVLVCLISDFIAGTGLGVLLSNAPI 69

```

Db 7 GLILGFSLLAIGSQNAFILRGILGRHVFLGCVNLSALLIVAGVSGFFVLVKKTPY 66
QY 70 VLDIWRWGIIAYLLWFAVNAKADMTNKVEAPQIIEETPTVDDTFLGGSVATDTR 129
Db 67 IVDVARYGGAFFILWYGI-----SRLLSALRGD 94
QY 130 VRVEVSVDKQVWV-KPMLMAVLTLWNPAYLDAFVFGVGAQYDGTGRWTFAGAGA 188
Db 95 SQMDISKAKEADSLPKALLSCIAFTLNPVHYLDTFVLGSGISAQFGDES-WKFGVGSA 153
QY 189 ASLWFLPVLGFGAALSRPL-SSPKVWRWVNVVAVMTALAIKMLM 235
Db 154 ASLIFFCGLGYGALL-RPLSSPVAVRILDFVIGVMTMLAGKLAPL 200

RESULT 2
Q8FE55 PRELIMINARY; PRT; 211 AA.
AC Q8FE55;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein yggA.
GN YGGA OR C3501.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=2238234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AB016766; AN81949.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005293; F:lysine permease activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro; IPR003016; Lipoyl_BS.
DR InterPro; IPR001123; Lyse.
DR InterPro; IPR004777; Lys_exporter.
DR Pfam; PF01810; Lyse; 1.
DR TIGRFAMs; TIGR00948; 2a75; 1.
DR PROSITE; PS00189; LIPOYL; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 211 AA; 23173 MW; 5F2D06A3A8FBD73 CRC64;

Query Match 29.1%; Score 347; DB 16; Length 211;
Best Local Similarity 35.7%; Pred. No. 2.6e-22;
Matches 81; Conservative 37; Mismatches 75; Indels 34; Gaps 4;

QY 7 FITGLIGALLSLTGPNVIVIKOGIKREGLIAVLVCLISDVFLIAGTLGVDLISNA 66
Db 5 YFQGLAIGAMILPLGPONAFVWQGIKRRQYHIMLALCALSDLVLCAGIFGGSALLMQ 64
QY 67 APVLDIMRWGGIAYLLWFAVNAKADMTNKVEAPQIIEETPTVDDTFLGGSVATDT 126
Db 65 SPWLLALVTWGGVAFLLWYGFAGFKTAMSSNIEL----- 98
QY 127 RNRVRVSVSDKQVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVW 185
Db 99 -----ASAELVKKQGRW-KIATMLAVTLNPNHYLDTFVLGSGQLDVFPEKRW-FALG 151
QY 186 AFAASLTWFLVCGFGAALSRPLSSPKVWRWVNVVAVMTALAIKL 232
Db 152 TISASFLWFGIALLAWLAPRLTKASQRIINLVVGCWVWFIALQL 198

```

```

RESULT 3
Q823W2 PRELIMINARY; PRT; 211 AA.
AC Q823W2;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Possible membrane transport protein.
GN STY3222 OR T2984.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
CX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RL Nature 413:848-852 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TY2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyourami V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
RL J. Bacteriol. 185:2330-2337 (2003).
DR EMBL; AL627277; CAD02896.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005293; F:lysine permease activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro; IPR001123; Lyse.
DR InterPro; IPR004777; Lys_exporter.
DR Pfam; PF01810; Lyse; 1.
DR TIGRFAMs; TIGR00948; 2a75; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 211 AA; 23172 MW; 13CB427CABESA3FC CRC64;

Query Match 29.1%; Score 347; DB 16; Length 211;
Best Local Similarity 33.6%; Pred. No. 2.6e-22;
Matches 79; Conservative 42; Mismatches 80; Indels 34; Gaps 4;

QY 3 IWEFIPTGLIGALLSLTGPNVIVIKOGIKREGLIAVLVCLISDVFLIAGTLGVDL 62
Db 1 MISTYFGVALGAMILPLGPONAFVWQGIKRRQYHIMLALCALSDLVLCAGIFGGS 60
QY 63 LSNAAPVLDIMRWGGIAYLLWFAVNAKADMTNKVEAPQIIEETPTVDDTFLGGS 122
Db 61 LLMOSPFLIALVTWGGVAFLLWYGFAGFKTAMSSNIEL----- 98
QY 123 ATDTRNRVRVSVSDKQVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVW 181
Db 99 -----ASAELVKKQGRW-KIATMLAVTLNPNHYLDTFVLGSGQLDVFPEKRW- 147
QY 182 FFAAGFAASLTWFLVCGFGAALSRPLSSPKVWRWVNVVAVMTALAIKMLMG 236
Db 148 FALGTISASFLWFGIALLAWLAPRLTKAQRIINLVVGCWVWVWFIALQALREG 202

RESULT 4
Q87X28

```





```

DR EMBL: AP002563; BAB37217.1; --
DR PIR: B91103; B91103.
DR PIR: B95948; B95948.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005293; F: lysine permease activity; IEA.
DR GO: GO:0006865; P: amino acid transport; IEA.
DR InterPro: IPR003016; Lipoyl_BS.
DR InterPro: IPR001123; LyseE.
DR InterPro: IPR004777; Lys_exporter.
DR Pfam: PF01810; LyseE; 1.
DR TIGRFAMs: TIGR00948; 2a75; 1.
DR PROSITE: PS00189; LIPOYL; 1.
KW Complete proteome.
SQ SEQUENCE 211 AA; 23202 MW; 4422D078B182266C CRC64;

Query Match 28.8%; Score 343; DB 16; Length 211;
Best Local Similarity 35.2%; Pred. No. 5.7e-22;
Matches 80; Conservative 37; Mismatches 76; Indels 34; Gaps 4;

QY 7 FITGLLGASLLSTGPONVLVIKQIKREGILAVLLVCLISDVFLFIAGTLGVLLSNA 66
DB 5 YFQGLAGAMILPLGPQNAFVNQGIHQYHIMIALICASDLVLCAGIFGSSALLMQ 64
QY 67 APIVLDIRMGGIAYILLNFANVAAKDMTNKVEAPQIIETEPTVPDDTPIGGSVAVD 126
DB 65 SPWLLALVTWGGVFWLLWYGFAGFKTAMSSNIEL----- 98

QY 127 RNRVVEVSDKORVWKPKMLMAVLTLNPNAYLDAFVIGVGGAQYG-DTGRWIFAAG 185
DB 99 -----ASAEVLKQGRW-KIATMTLAVTLNPNHVIDLTFVVLGSLGQDLVDPKRW-FALG 151

QY 186 AFAASLIWFLPVGFAAALSRPLSSPKWWRINVVAVVMTALAIXL 232
DB 152 TISASFLWFFGLAILAIAWAPRLRTAKSORIINLVGCVWVFALQL 198

RESULT 7
QY QYUBP8 PRELIMINARY; PRT; 211 AA.
AC QYUBP8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Hypothetical protein yggA.
GN YGG A OR S3108.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OC NCBI_TaxID=623;
RN RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL: AE016988; AAP18212.1; --
KW Hypothetical protein.
SQ SEQUENCE 211 AA; 23201 MW; 4422D078B182266C CRC64;

Query Match 28.8%; Score 343; DB 16; Length 211;
Best Local Similarity 35.2%; Pred. No. 5.7e-22;
Matches 80; Conservative 37; Mismatches 76; Indels 34; Gaps 4;

QY 7 FITGLLGASLLSTGPONVLVIKQIKREGILAVLLVCLISDVFLFIAGTLGVLLSNA 66
DB 5 YFQGLAGAMILPLGPQNAFVNQGIHQYHIMIALICASDLVLCAGIFGSSALLMQ 64
QY 67 APIVLDIRMGGIAYILLNFANVAAKDMTNKVEAPQIIETEPTVPDDTPIGGSVAVD 126
DB 65 SPWLLALVTWGGVFWLLWYGFAGFKTAMSSNIEL----- 98

QY 127 RNRVVEVSDKORVWKPKMLMAVLTLNPNAYLDAFVIGVGGAQYG-DTGRWIFAAG 185
DB 99 -----ASAEVLKQGRW-KIATMTLAVTLNPNHVIDLTFVVLGSLGQDLVDPKRW-FALG 151

QY 186 AFAASLIWFLPVGFAAALSRPLSSPKWWRINVVAVVMTALAIXL 232
DB 152 TISASFLWFFGLAILAIAWAPRLRTAKSORIINLVGCVWVFALQL 198

RESULT 7
QY QYUBP8 PRELIMINARY; PRT; 211 AA.
AC QYUBP8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Hypothetical protein yggA.
GN YGG A OR S3108.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OC NCBI_TaxID=623;
RN RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL: AE016988; AAP18212.1; --
KW Hypothetical protein.
SQ SEQUENCE 211 AA; 23201 MW; 4422D078B182266C CRC64;

Query Match 28.8%; Score 336; DB 16; Length 212;
Best Local Similarity 31.0%; Pred. No. 2.3e-21;
Matches 71; Conservative 44; Mismatches 82; Indels 32; Gaps 1;

QY 8 ITGLLGASLLSTGPONVLVIKQIKREGILAVLLVCLISDVFLFIAGTLGVLLSNA 67
DB 16 LSGPGLGASLIVAIGNAQNAVYVLRQGLREYVVGVLICALCDMALIALGVAGMGTLSAH 75
QY 68 PIVLDIMRWGGIAYILLNFANVAAKDMTNKVEAPQIIETEPTVPDDTPIGGSVAVD 127
DB 76 PAULTAVRWAGAFLLAYGARAFRAW-----R 103
QY 128 RNRVVEVSDKORVWKPKMLMAVLTLNPNAYLDAFVIGVGGAQYDGTGRWIFAAG 187
DB 104 GAERLQARNGDKASHAQVLAASLILNPHVYDVTVLLGAGRYAMPANVAFAGAM 163
QY 188 AASLIWFLPVGFAAALSRPLSSPKWWRINVVAVVMTALAIXL 236
DB 164 CASILWFSLLGFGARLLEPFAFVAPVAVRVLDAIGAVMAIALTLLMG 212

RESULT 9
QY Q8ZHH6 PRELIMINARY; PRT; 205 AA.
AC Q8ZHH6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

```

```

DB 65 SPWLLALVTWGGVFWLLWYGFAGFKTAMSSNIEL----- 98

QY 127 RNRVVEVSDKORVWKPKMLMAVLTLNPNAYLDAFVIGVGGAQYG-DTGRWIFAAG 185
DB 99 -----ASAEVLKQGRW-KIATMTLAVTLNPNHVIDLTFVVLGSLGQDLVDPKRW-FALG 151

QY 186 AFAASLIWFLPVGFAAALSRPLSSPKWWRINVVAVVMTALAIXL 232
DB 152 TISASFLWFFGLAILAIAWAPRLRTAKSORIINLVGCVWVFALQL 198

RESULT 8
QY Q8Y2F1 PRELIMINARY; PRT; 212 AA.
AC Q8Y2F1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Probable transmembrane protein.
GN RSC0385 OR RSC03353.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OC NCBI_TaxID=305;
RN RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Attiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cumac S., Demange N.,
RA Gaspin C., Lavi M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502 (2002).
DR EMBL: AL646059; CAD13913.1; --
DR GO: GO:0016020; C: membrane; IEA.
DR GO: GO:0005293; F: lysine permease activity; IEA.
DR GO: GO:0006865; P: amino acid transport; IEA.
DR InterPro: IPR001123; LyseE.
DR Pfam: PF01810; LyseE; 1.
KW Complete proteome.
SQ SEQUENCE 212 AA; 21896 MW; E29D663C9497BDBB CRC64;

Query Match 28.2%; Score 336; DB 16; Length 212;
Best Local Similarity 31.0%; Pred. No. 2.3e-21;
Matches 71; Conservative 44; Mismatches 82; Indels 32; Gaps 1;

QY 8 ITGLLGASLLSTGPONVLVIKQIKREGILAVLLVCLISDVFLFIAGTLGVLLSNA 67
DB 16 LSGPGLGASLIVAIGNAQNAVYVLRQGLREYVVGVLICALCDMALIALGVAGMGTLSAH 75
QY 68 PIVLDIMRWGGIAYILLNFANVAAKDMTNKVEAPQIIETEPTVPDDTPIGGSVAVD 127
DB 76 PAULTAVRWAGAFLLAYGARAFRAW-----R 103
QY 128 RNRVVEVSDKORVWKPKMLMAVLTLNPNAYLDAFVIGVGGAQYDGTGRWIFAAG 187
DB 104 GAERLQARNGDKASHAQVLAASLILNPHVYDVTVLLGAGRYAMPANVAFAGAM 163
QY 188 AASLIWFLPVGFAAALSRPLSSPKWWRINVVAVVMTALAIXL 236
DB 164 CASILWFSLLGFGARLLEPFAFVAPVAVRVLDAIGAVMAIALTLLMG 212

RESULT 9
QY Q8ZHH6 PRELIMINARY; PRT; 205 AA.
AC Q8ZHH6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

```

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Putative LysE type translocator (Hypothetical protein).  
GN YP00918 OR Y305.  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / Biovar Orientalis;  
RX MEDLINE=21470413; PubMed=11586360;  
RA Parkhill J., Wren B.W., Thomson N.R., Tittball R.W., Holden M.T.G.,  
RA Partridge M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltwell T., Hamlin N., Holtroyd S., Jagels K., Karyshev A.V.,  
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;  
RT "Genome sequence of Yersinia pestis, the causative agent of plague."  
RL Nature 413:523-527(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KIM5 / Biovar Mediaevalis;  
RX MEDLINE=22137863; PubMed=12142430;  
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
RA Perry R.D.;  
RT "Genome sequence of Yersinia pestis KIM."  
RL J. Bacteriol. 184:4601-4611(2002).  
DR EMBL: AJ414145; CAC89762.1; -;  
DR EMBL: AE013931; AAM6855.1; -;  
DR PIR: AG0112; AG0112.  
DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0005293; F: lysine permease activity; IEA.  
DR GO: GO:0006865; P: amino acid transport; IEA.  
DR InterPro: IPR001123; LysE.  
DR InterPro: IPR004777; Lys\_exporter.  
DR Pfam: PF01810; LysE; 1.  
DR TIGRfam: TIGR00948; 2a75; 1.  
DR Hypothetical protein; Complete proteome.  
KW SEQUENCE 205 AA; 22164 MW; 2969524391474CDA CRC64;  
Query Match 28.1%; Score 335; DB 16; Length 205;  
Best Local Similarity 33.1%; Pred. No. 2.7e-21;  
Matches 80; Conservative 44; Mismatches 70; Indels 48; Gaps 6;  
QY 3 IMEIFTGLLGSALLSTGPQNVIVIKQIKREGIIVLVCLISDVFLFIAGTLGVDL 62  
DB 1 MLAVYIHGFILSAAMILPLQPNQVFMVNOGKIKRQHLMSASICALSDIILICAGIFGSA 60  
QY 63 LSNAPVILDTMRGGIAYLLWF---AVMAAKDAMTKVEAPQIIEETPTVDDTPLGG 119  
DB 61 LLSRSPLLALVTWGGVAFIMVYGVGALMAA-----LAVTLNPHVYLDTFVVLGSLGGQL 141  
QY 120 SAVATD----TNRVRVSVSKQKRWVVKPMLMAIVLTWLNPNAYLDAPFVIGGGAQ-Y 174  
DB 95 DGVASATSVTQGRWRLVTL-----LAVTLNPHVYLDTFVVLGSLGGQL 141  
QY 175 GDTGRWIFNAGAFASLINFPLVFGGAALSRPLSSPKVWRWVNVVAVMTALAKML 234  
DB 142 PDIRPW-FALGAVTASIVWFVAFALALAAWLSFMINRPVQRIINIFVGGVWGFIAPOLAR 200  
QY 235 MG 236  
DB 201 QG 202  
RESULT 10  
ID Q9HW36 PRELIMINARY; PRT; 200 AA.  
AC Q9HW36;

DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Probable transporter.  
GN PA4365.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Hickey M.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Garber R.L., Goltzy S.N., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen."  
RL Nature 406:959-964(2000).  
DR EMBL: AE004852; AAG07753.1; -;  
DR PIR: D83100; D83100.  
DR GO: GO:0016020; C: membrane; IEA.  
DR GO: GO:0005489; F: electron transporter activity; IEA.  
DR GO: GO:0005293; F: lysine permease activity; IEA.  
DR GO: GO:0006865; P: amino acid transport; IEA.  
DR GO: GO:0006118; P: electron transport; IEA.  
DR InterPro: IPR002109; Glutaredoxin.  
DR InterPro: IPR001123; LysE.  
DR Pfam: PF01810; LysE; 1.  
DR PROSITE: PS00195; GLUTAREDOXIN; 1.  
KW Complete proteome.  
SQ SEQUENCE 200 AA; 21144 MW; 7C10CB8D8FE36A4 CRC64;  
Query Match 27.9%; Score 332; DB 16; Length 200;  
Best Local Similarity 33.5%; Pred. No. 4.8e-21;  
Matches 77; Conservative 42; Mismatches 77; Indels 34; Gaps 4;  
QY 7 FITGLLGSALLSTGPQNVIVIKQIKREGIIVLVCLISDVFLFIAGTLGVDLSNA 66  
DB 5 YLNGIIVAAGLIAGIAGNAFVLAQSLRREHLSVAALCVFCDVIVSLGVFLGLKILLE 64  
QY 67 APIVLDIRWGGIAYLLWFVAVMAAKDAMTKVEAPQIIEETPTVDDTPLGGSAVATDT 126  
DB 106 RSRKAV-----LAAALAVTLNPHVYLDTFVVLGSLGGAQAAPG--AYALGA 150  
QY 187 FAASLWFFLVFGGAALSRPLSSPKVWRWVNVVAVMTALAKMLMG 236  
DB 151 ASASLWFFVAFALAGVAFVLAQSLRREHLSVAALCVFCDVIVSLGVFLGLKILLE 200  
RESULT 11  
ID Q9K4K6 PRELIMINARY; PRT; 204 AA.  
AC Q9K4K6  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Putative membrane transport protein.  
GN SC07308 OR SC5F8.18.  
OS Streptomyces coelicolor.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);

**Seeger K.J., Harris D.;**

Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.

[2]

RN  
RN  
RP  
SEQUENCE FROM N.A.  
STRAIN=A3(2);  
Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.

[3]

RN  
RN  
RP  
SEQUENCE FROM N.A.  
STRAIN=A3(2);  
MEDLINE=97000351; PubMed=8843436;  
Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,  
Kinashi H., Hopwood D.A.;  
RNA set of ordered cosmids and a detailed generic and physical map for  
the 8 MB Streptomyces coelicolor A3(2) chromosome.";

Mol. Microbiol. 21:77-96 (1996).

[4]

RN  
RN  
RP  
SEQUENCE FROM N.A.  
STRAIN=A3(2) / MI45;  
MEDLINE=21996410; PubMed=12000953;  
Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
Cronin A., Fraser A., Goale A., Hidalgo J., Hornsbey T., Howarth S.,  
Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,  
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,  
Hopwood D.A.);  
RT "Complete genome sequence of the model actinomycete Streptomyces  
coelicolor A3(2).";  
RL Nature 417:141-147(2002).  
DR EMBL; AL939131; CAB93746.1; -.  
GO; GO:0016020; C:membrane; IEA.  
DR DR GO; GO:0005293; F:lysine permease activity; IEA.  
DR GO; GO:0006865; P:amino acid transport; IEA.  
DR InterPro; IPR001123; LyseE.  
DR Pfam; PF01810; LyseI; 1.  
KW Complete proteome.  
SQ SEQUENCE 204 AA; 20686 MW; DD1BD686CC04DA1 CRC64;

Query Match 27.6%; Score 328.5; DB 16; Length 204;  
Best Local Similarity 34.8%; Pred. No. le-20; Indels 33; Gaps 4;  
Matches 78; Conservative 32; Mismatches 81;

QY 10 GLLLGASILLISGPNNVLVIKGIKRSGIHAYLVLCITSDVFLFIAGTLGVCLLSNAPI 69  
Db :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Dy 11 GFCTGLSLVAIGAQNAFVLRGVRDLAVLVGGICALSADVILALGVGGVVWAPG 70  
  
QY 70 VLDIMRWGGIIAYLIIFANVAADKANVTKEAQITEETFTVPDDTPLCGSSAVATDNRR 129  
Db ||:||:||||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Dy 71 ALTAVGWGTGGAFLLCYGAIAAR-----RVFRP-----SGAIRADAAGAARRR 113  
  
QY 130 VRVEVSVDQRVMVKPMMLMAIVLTWLNPNIAYDAVFPTIGVGQAQYGDTGRWFPAAGAAFA 189  
Db ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Dy 114 A-----VVICIALTWLNPHYLDTFVLGSVAADRGPL-RWTFGLGAAA 157  
  
QY 190 SLIWPLPVGFCAALSRLPSPKVWRNVINNVAVVMTALAIKLM 233  
Db |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Dy 158 SWLFWAALGFGARLYGRFLSRVNAVNRVIDGLVNATVILVIGSVLV 201  
  
RESULT 12  
QBHW5 PRELIMINARY; PRt; 208 AA.  
ID QBHW5 AC QBHW5;  
AC QBHW5;  
DT 01-OCT-2002 (TEMBELrel. 22, Created)  
DT 01-OCT-2002 (TEMBELrel. 22, Last sequence update)  
DT 01-JUN-2003 (TEMBELrel. 24, Last annotation update)  
DE Membrane transport protein.  
GN YPGA OR XAC3133.  
OS Xanthomonas axonopodis (pv. citri).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OT Xanthomonas axonopodis (pv. citri).

RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kufayin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*  
 RT C58.",  
 RT Science 294:2317-2323 (2001).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT *Agrobacterium tumefaciens* C58.",  
 RT Science 294:2323-2328 (2001).  
 RL EMBL; AB009057; RA141941.1; -;  
 DR PIR; AG2690; AG2690.  
 DR PIR; C97472; C97472.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005293; F:lysine permease activity; IEA.  
 DR GO; GO:0006865; P:amino acid transport; IEA.  
 DR InterPro; IPR001123; LysE.  
 DR Pfam; PF01810; LysE; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 202 AA; 21943 MW; F43E0FF2A834D16 CRC64;

Query Match 27.4%; Score 326.5; DB 16; Length 202;  
 Best Local Similarity 31.6%; Pred. No. 1.5e-20;  
 Matches 73; Conservative 42; Mismatches 83; Indels 33; Gaps 4;

QY 4 MEIFITGLGASLLLSIGPQNVLIKQIKREGIAVLVCLISDVFLFIAGTGLVDLL 63  
 DB 3 IQIFFTGLTGLSLVIAIGQNAFLVKGARSHVAVCATCAISDALLWGVGFQRI 62

QY 64 SNAAPITVLDIMRWGGIAYLLWFVAVMAKAMTKNKEAPQIIEETPTVDDTPLGSAVA 123  
 DB 63 SAIMPALDIPMKYAGAAFLIYWGAKSLYALRSS-EVLSVAERREAS----- 108

QY 124 TDRNRVRVSVQKRVVWKPMLMAIVLTWLNPNAYLDAFVFGVCAQYGTGWFIFA 183  
 DB 109 -----LW-OTLAICLALTFLNPHVYLDVTLGTISTQPGFEK-TFA 149

QY 184 AGAFAASLWFLPVGGAALSRPLSSPKVWRVWVAVVMTALAIKML 234  
 DB 150 AGAATOSLLFFSLGYGARWLRDIPEKPSANRILEGVIALTWATFALVW 200

RESULT 14  
 Q88PD4 PRELIMINARY; PRT; 204 AA.  
 AC Q88PD4; (T-EMBLrel. 24, Created)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)  
 DE 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE Transporter, LysE family.  
 GN PP0916.  
 OS *Pseudomonas putida* (strain KT2440).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=160488;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=22423060; PubMed=12534463;  
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,

RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,  
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,  
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,  
 RA Utterback T., Rizzo M., Lee K., Kosack D., Moesti D., Wedler H.,  
 RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,  
 RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoft A., Tummeler B.,  
 RA Frazer C.M.;  
 RT "Complete genome sequence and comparative analysis of the  
 RT metabolically versatile *Pseudomonas putida* KT2440.",  
 RL Environ. Microbiol. 4:799-808 (2002).  
 DR EMBL; AB016777; AAN66541.1; -;  
 DR TIGR; PP0916; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005293; F:lysine permease activity; IEA.  
 DR GO; GO:0006865; P:amino acid transport; IEA.  
 DR InterPro; IPR001123; LysE.  
 DR Pfam; PF01810; LysE; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 204 AA; 21665 MW; 724CFB736E2B12F CRC64;

Query Match 27.1%; Score 323; DB 16; Length 204;  
 Best Local Similarity 32.2%; Pred. No. 3e-20;  
 Matches 73; Conservative 47; Mismatches 73; Indels 34; Gaps 5;

QY 7 FITGLLGLSLLLSIGPQNVLIKQIKREGIAVLVCLISDVFLFIAGTGLVDLLSNA 66  
 DB 10 YLNGMLVAFGLMAIGAQNAFLVLAQSLRREHLPVAALCIVCDAILVAAGVGLATVLAH 69

QY 67 APIVLDIMRWGGIAYLLWFVAVMAKAMTKNKEAPQIIEETPTVDDTPLGSSAVATDT 126  
 DB 70 NPTLLAIARWGGVAVFLIYWGAKSLYALRSS-----QSLQHQ-----GGGV----- 110

QY 127 RNRVRVSVQKRVVWKPMLMAIVLTWLNPNAYLDAFVFGVCAQYGTGWRIFAAGA 186  
 DB 111 RSRRAV-----LLSALVTLNPHVYLDVTLGLSGLGQSQSAGVAV--AGA 155

QY 187 PAASLWFLPVGGAALSRPLSSPKVWRVWVAVVMTALAIKML 233  
 DB 156 ASASLLWFLSTLAIGAALPWLARPATWRLDLMVAVMFAVAQLI 202

RESULT 15  
 Q8P6L3 PRELIMINARY; PRT; 208 AA.  
 AC Q8P6L3; (T-EMBLrel. 22, Created)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DE 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
 DE Membrane transport protein.  
 GN YGGA OR KCC2954.  
 OS *Xanthomonas campestris* (pv. campestris).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OX NCBI\_TaxID=340;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=ATCC 33913 / NCPPB 528;  
 RC MEDLINE=22022145; PubMed=12024217;  
 RX da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Qaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A., P.,  
 RA Cicarotte G., Camavari F., Cardozo J., Chamargo F., Clapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kitajima J.P.;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:15:20 ; Search time 9505 Seconds  
(without alignments)  
10825.501 Million cell updates/sec

Title: US-09-105-117K-3  
Perfect score: 2374  
Sequence: 1 agatactcttgaagaaa.....gtacacattcagcaaatgg 2374

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pac.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sw.\*
- 13: gb\_un.\*
- 14: gb\_vt.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_ma.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_hgt\_hum.\*
- 31: em\_hgt\_inv.\*
- 32: em\_hgt\_ther.\*
- 33: em\_hgt\_mus.\*
- 34: em\_hgt\_pln.\*
- 35: em\_hgt\_rod.\*
- 36: em\_hgt\_mam.\*
- 37: em\_hgt\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hgt\_hum.\*
- 40: em\_hgt\_mus.\*
- 41: em\_hgt\_ther.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	2374	100.0	2374	1	CGLYSEG	X96471 C-glutaminc
C 2	2374	100.0	2374	6	A93933	A93933 Sequence 2
C 3	2374	100.0	333150	1	AP005277	AP005277 Coryneb
C 4	2374	100.0	349980	6	AX127147	AX127147 Sequence
C 5	993	41.8	993	6	AX063767	AX063767 Sequence
C 6	993	41.8	993	6	AX244055	AX244055 Sequence
C 7	870	36.6	870	6	AX123540	AX123540 Sequence
C 8	870	36.6	870	6	BD165657	BD165657 Novel pol
C 9	822	34.6	822	6	AX063771	AX063771 Sequence
C 10	822	34.6	822	6	AX244059	AX244059 Sequence
C 11	818.6	34.5	308650	1	AB005218	AB005218 Coryneb
C 12	798.8	33.6	1771	1	AB083133	AB083133 Coryneb
C 13	724.4	30.5	349535	1	BX248357	BX248357 Coryneb
C 14	711	29.9	711	6	AX643028	AX643028 Sequence
C 15	708	29.8	708	6	AX123539	AX123539 Sequence
C 16	708	29.8	708	6	BD165656	BD165656 Novel pol
C 17	698.6	29.4	1568	6	E54483	E54483 Heat-resist
C 18	696.8	29.4	1083	6	AX813971	AX813971 Sequence
C 19	649	27.3	1083	6	AX123538	AX123538 Sequence
C 20	627	26.4	627	6	BD165655	BD165655 Novel pol
C 21	627	26.4	627	6	AX067087	AX067087 Sequence
C 22	485.2	20.4	993	6	AP003001	AP003001 Mesorhizo
C 23	170	7.2	345783	1	AP003001	AP003001 Mesorhizo
C 24	164.4	6.9	8472	1	AE008076	AE008076 Agrobacte
C 25	164.4	6.9	13051	1	AE009111	AE009111 Agrobacte
C 26	159.4	6.7	302007	1	SC0939132	SC0939132 Streptomy
C 27	140.6	5.9	303550	1	SC0939131	SC0939131 Streptomy
C 28	135.6	5.7	306250	1	SMES591788	SMES591788 Sinorhizo
C 29	123.8	5.2	347356	1	BX640437	BX640437 Bordetell
C 30	123.8	5.2	348171	1	BX640412	BX640412 Bordetell
C 31	122.8	5.2	14844	1	MBU34849	MBU34849 Mycobacteri
C 32	122.8	5.2	15239	6	I86263	I86263 Sequence 17
C 33	122.8	5.2	306050	1	BX248341	BX248341 Mycobacte
C 34	122.2	5.1	348251	1	BX640423	BX640423 Bordetell
C 35	121.2	5.1	14869	1	AE007056	AE007056 Mycobacte
C 36	121.2	5.1	346186	1	BX842578	BX842578 Mycobacte
C 37	120.8	5.1	298300	1	AP005025	AP005025 Streptomy
C 38	116.4	4.9	11120	1	FEA311775	FEA311775 Rhodococc
C 39	115.2	4.9	298300	1	AP005025	AP005025 Streptomy
C 40	113	4.8	303550	1	SC0939131	SC0939131 Streptomy
C 41	112.6	4.7	10871	1	AE011965	AE011965 Xanthomon
C 42	112.4	4.7	10336	1	AE005530	AE005530 Escherich
C 43	112.2	4.7	12070	1	AB011413	AB011413 Streptomy
C 44	111.2	4.7	10278	1	AE015317	AE015317 Shigella
C 45	111.2	4.7	266658	1	AP002563	AP002563 Escherich

ALIGNMENTS

RESULT 1  
CGLYSEG/c  
LOCUS  
DEFINITION C-glutaminc lysE and lysG genes.  
ACCESSION X96471  
VERSION X96471.1 GI:1729753  
KEYWORDS lysE gene; lysG gene; lysine export regulator protein; lysine exporter protein; lysine governor.  
SOURCE Corynebacterium glutamicum  
ORGANISM Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
REFERENCE 1  
AUTHORS Vrljic, M., Salm, H. and Eggeling, L.  
TITLE A new type of transporter with a new type of cellular function:

JOURNAL: l-lysine export from *Corynebacterium glutamicum*  
MEDLINE: Mol. Microbiol. 22 (5), 815-826 (1996)  
PUBMED: 89716810  
AUTHORS: 2 (bases 1 to 2374)  
REFERENCE: Vrljic, M.M.  
TITLE: Direct Submission  
JOURNAL: Submitted (07-MAR-1996) M.M. Vrljic, Institut fuer Biotechnologie  
1. Forschungszentrum Juelich, Postfach 1913, D-52425 Juelich, FRG  
FEATURES  
source  
1. ..2374  
/organism="Corynebacterium glutamicum"  
/mol\_type="genomic DNA"  
/strain="R127"  
/db\_xref="taxon:1718"  
complement (82. .954)  
/gene="lysG"  
complement (82. .954)  
/gene="lysG"  
/note="lysine governor"  
/codon\_start=1  
/transl\_table=11  
/product="lysine export regulator protein"  
/protein\_id="CAA65323.1"  
/db\_xref="GI:1729754"  
/db\_xref="GOA:P94632"  
/db\_xref="SWISS-PROT:P94632"  
/translation="MNPIQLDTLLSIIDEGSPGASLALSTSPSAVSORVKALRHVVG  
RVLSRQPAKATEAGEVLVQAKRWLLQASTRAQISGRLEAIPITIAINADSLSTW  
PPFVFNVSAGGATLTLRDEAHTLSLRGDLVIGAVTREPANPVAGCEVVELGTMR  
HLATATSLDRAVDWGLKDWAMPVLRFGPKDVLQDRDLGRVDFGVRRRVISVPS  
ABGFGEARLRLGWLGPETQAPMLKAGEVILLDEIPIDTPMYQWRWLESRLAEL  
TDAVVDAIIEGLRP"  
1025. .1726  
/gene="lysE"  
1025. .1726  
/gene="lysE"  
/codon\_start=1  
/transl\_table=11  
/product="Lysine exporter protein"  
/protein\_id="CAA65324.2"  
/db\_xref="GI:13397387"  
/db\_xref="GOA:P94633"  
/db\_xref="SWISS-PROT:P94633"  
/translation="MEIFITGLIGASLLISHPQNVILVKKQIKREGILAVLLVCLH  
SDVFLTAGTLGVLLSNAPITVLDINRWGGLAYLLWFAVMAKADAMTKVAPQIIE  
ETEPVDDPLGSSAVATDRNEVRVENSVDKORVWVKPMLMAIIVLWLNPNVLA  
FVFGGCAQYGDGTGRWIFPAGAPAAASLIWFLVGFVGAALSRPLSSPKWRINVVV  
AVVMTALAILMLWG"  
complement (1723. .2352)  
/gene="orf3"  
complement (1723. .2352)  
/gene="orf3"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA65325.1"  
/db\_xref="GI:1729756"  
/db\_xref="SPTREMBL:P94634"  
/translation="WVALRDIIVASKALYVGISSYGPETLTAABEFMAWREGCPILLHQ  
PYSIINRWEEPPDDGNNLQSANNGLVIAFSPFLAQILLTKYLDGIPGSRASQ  
GKSLSEGLMNVNLDWRKINDIAGRGSLAQMALWLRQEGYGDVITVSALIGA  
SSVEQLDNLSDLSNLEFSDAELEAIDEISHDGINIWAKATDSKTREN"  
ORIGIN  
Query Match 100.0%; Score 2374; DB 1; Length 2374;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGATACCTCTTTGGAGAAACCATGTACCGATTGCGTACATTTGTTGGTCTGGAAGGC 60  
Db 2374 AGATACCTCTTTGGAGAAACCATGTACCGATTGCGTACATTTGTTGGTCTGGAAGGC 2315  
Qy 61 TCTTTAGCTGGGTATTCTTCTTACGGTCCAGAGCTCACAGCGAGGCGGCTGAGTTCAT 120

2314 TCTTTAGCTGGGTATTCTTCTTACGGTCCAGAGCTCACAGCGAGGCGGCTGAGTTCAT 2255  
Qy 121 GCGGAGAGGGCTGCGCGCTTCTGATTATCATCAGCAGCAAGCTATTCATCATTAATCGTGG 180  
Db 2254 GCGGAGAGGGCTGCGCGCTTCTGATTATCATCAGCAGCAAGCTATTCATCATTAATCGTGG 2195  
Qy 181 GGTGAGGAACCGGGCGATGACGCTGAGAACTTGTGTGCACTGAGCTGCGCAACAATGGTCT 240  
Db 2194 GGTGAGGAACCGGGCGATGACGCTGAGAACTTGTGTGCACTGAGCTGCGCAACAATGGTCT 2135  
Qy 241 TGGGCTCATGTGTTTCTACCATTTGCGCAGGGCTGTCTCACGGAACAATATCTCGATGG 300  
Db 2134 TGGGCTCATGTGTTTCTACCATTTGCGCAGGGCTGTCTCACGGAACAATATCTCGATGG 2075  
Qy 301 AATTCCAGAGGGTTCCCGCGCAGCAGGTAAGTCCCTGTCTGAGGGGCATGTTGAACGT 360  
Db 2074 AATTCCAGAGGGTTCCCGCGCAGCAGGTAAGTCCCTGTCTGAGGGGCATGTTGAACGT 2015  
Qy 361 GAACAATATTGATATGTTGTCGCAAGCTCAATGACATGCCCGCAGGAACCGCGGCACTCACT 420  
Db 2014 GAACAATATTGATATGTTGTCGCAAGCTCAATGACATGCCCGCAGGAACCGCGGCACTCACT 1955  
Qy 421 TGGCGAGATGGGCTTGCATGGTGTCTGGCGCAGCAAGAGATGACGGCGGATACCGT 480  
Db 1954 TGGCGAGATGGGCTTGCATGGTGTCTGGCGCAGCAAGAGATGACGGCGGATACCGT 1895  
Qy 481 GACGAGTGCATTGATTGGTCTTGTGCTGAGTTCAGCTGAGCAACAGCTTCAATCACT 540  
Db 1894 GACGAGTGCATTGATTGGTCTTGTGCTGAGTTCAGCTGAGCAACAGCTTCAATCACT 1835  
Qy 541 CAACAATTTGGAGTTTCTGACGCCGAGTGGAGCGATGATGATGATTTCCCAACGAGCG 600  
Db 1834 CAACAATTTGGAGTTTCTGACGCCGAGTGGAGCGATGATGATGATTTCCCAACGAGCG 1775  
Qy 601 CGGCTCAACATTTGGGCGAAGCCACCGATTCCAAACCGCGCAAACTAACCCATCAA 660  
Db 1774 CGGCTCAACATTTGGGCGAAGCCACCGATTCCAAACCGCGCAAACTAACCCATCAA 1715  
Qy 661 CATCAGTTTGTATGGCCAAATCGGTCAATCACACTGCCACGACGAGTTCAGTCCAGCGCCA 720  
Db 1714 CATCAGTTTGTATGGCCAAATCGGTCAATCACACTGCCACGACGAGTTCAGTCCAGCGCCA 1655  
Qy 721 CACCTTGGGCTTGGACAGCGGGCGTGACAACTGTCTGTGGCGGAACCCACGCGGGA 780  
Db 1654 CACCTTGGGCTTGGACAGCGGGCGTGACAACTGTCTGTGGCGGAACCCACGCGGGA 1595  
Qy 781 CCAGTACAGGCTTGGCGGAAACCGCGGCGGAAATCCACGCTCGGCTGTCCCGGTA 840  
Db 1594 CCAGTACAGGCTTGGCGGAAACCGCGGCGGCGGAAATCCACGCTCGGCTGTCCCGGTA 1535  
Qy 841 TTGCGCGCGCAGCGCGCGGATAAACACAAACGCGTCCAAATAGCATTCGGGTTCAACCA 900  
Db 1534 TTGCGCGCGCAGCGCGCGGATAAACACAAACGCGTCCAAATAGCATTCGGGTTCAACCA 1475  
Qy 901 GGTGAGCAGATGGCCATCAACATGGGCTTTTACCAAAACCGCTGCTTATCGAGCTCAC 960  
Db 1474 GGTGAGCAGATGGCCATCAACATGGGCTTTTACCAAAACCGCTGCTTATCGAGCTCAC 1415  
Qy 961 CTCACACCGCAGCGGTTGGCGTGTGAGTGGCCACCGCGCAACCGCGCAAGCGGTGTC 1020  
Db 1414 CTCACACCGCAGCGGTTGGCGTGTGAGTGGCCACCGCGCAACCGCGCAAGCGGTGTC 1355  
Qy 1021 ATCGGCGCAGGTTGGTCTGTTTCTTCAATGATCTGTGGCGCTTCCACTTGTGTTTCAT 1080  
Db 1354 ATCGGCGCAGGTTGGTCTGTTTCTTCAATGATCTGTGGCGCTTCCACTTGTGTTTCAT 1295  
Qy 1081 GGGCTCTTTCGCTGCATGAGCGCAACCAATAGAGTATGCCATGCCAGCGCAT 1140  
Db 1294 GGGCTCTTTCGCTGCATGAGCGCAACCAATAGAGTATGCCATGCCAGCGCAT 1235  
Qy 1141 AATATCGAGCAGGATGGCGCGCATTTGAGCAAAAGATCAACCGCCCAAGGTGCGCGCAT 1200



Qy	2281	ATTGCGGCTTATGTTACTTCTGAAAGAGTTCAGGTTTTCATCTTCTGCGCCGCGAGAA	2340
Db	94	ATTGCGGCTTATGTTACTTCTGAAAGAGTTCAGGTTTTCATCTTCTGCGCCGCGAGAA	35
Qy	2341	TTGGGCCAGGCAGGTAACACCTTCAGCAAAATGG	2374
Db	34	TTGGGCCAGGCAGGTAACACCTTCAGCAAAATGG	1
RESULT 2			
A93933/c			
LOCUS	A93933	2374 bp	DNA linear PAT 22-JAN-2000
DEFINITION	Sequence 2 from Patent WO9723597.		
ACCESSION	A93933		
VERSION	A93933.1		
KEYWORDS	GI:6742037		
SOURCE	Corynebacterium glutamicum		
ORGANISM	Corynebacterium glutamicum		
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
AUTHORS	Corynebacterineae; Corynebacteriaceae; Corynebacterium.		
TITLE	1 (bases 1 to 2374)		
JOURNAL	Vrljic,M. and Egeling,L.		
FEATURES	PROCESS FOR THE MICROBIAL PRODUCTION OF AMINO ACIDS BY BOOSTED		
source	ACTIVITY OF EXPORT CARRIERS		
	Patent: WO 9723597-A 2 03-JUL-1997;		
	KERNFORSCHUNGSANLAGE JUELICH (DE); VRLJIC MARINA (DE)		
	Location/Qualifiers		
	1..2374		
	/organism="Corynebacterium glutamicum"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:1718"		
ORIGIN			
	Query Match	100.0%;	Score 2374; DB 6; Length 2374;
	Best Local Similarity	100.0%;	Pred. No. 0;
	Matches 2374;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	AGATACTCCTTTGGAGAAAACATGTACGATTCGGTGGA	ATTGTTGGTCTGGAAGGC 60
Db	2374	AGATACTCCTTTGGAGAAAACATGTACGATTCGGTGGA	ATTGTTGGTCTGGAAGGC 2315
Qy	61	TCCTTTACGTGGGTATTTCTTCTACGGTCCAGAGCTC	CACAGCGGAGGGCGGTGATTCAT 120
Db	2314	TCCTTTACGTGGGTATTTCTTCTACGGTCCAGAGCTC	CACAGCGGAGGGCGGTGATTCAT 2255
Qy	121	GGCGAGGAGGGCTCGCCGCTTCTGATTCATCAGCAAG	CTATTCATCATTAATCGTTG 180
Db	2254	GGCGAGGAGGGCTCGCCGCTTCTGATTCATCAGCAAG	CTATTCATCATTAATCGTTG 2195
Qy	181	GGTGAGGACCGGGCGATGAGGTGAGAACTTGTTCAGT	CAGTTCAGTCCACAAATCGTCT 240
Db	2194	GGTGAGGACCGGGCGATGAGGTGAGAACTTGTTCAGT	CAGTTCAGTCCACAAATCGTCT 2135
Qy	241	TGCGGTCAATTGCTTTCTCACCACTTGCGGACGGGCT	CTCCATCAGCAAAATATCTCCATGG 300
Db	2134	TGCGGTCAATTGCTTTCTCACCACTTGCGGACGGGCT	CTCCATCAGCAAAATATCTCCATGG 2075
Qy	301	AATTCAGAGGGTTCGCGCGCAGCAGAGGTAAGTCCCT	CTGTGAGGCGCATGTGAAAGT 360
Db	2074	AATTCAGAGGGTTCGCGCGCAGCAGAGGTAAGTCCCT	CTGTGAGGCGCATGTGAAAGT 2015
Qy	361	GAACAATATGATATGGTCCGCAAGCTGTAATGACAT	CGCCCGAGGAACGGCGCAGTCACT 420
Db	2014	GAACAATATGATATGGTCCGCAAGCTGTAATGACAT	CGCCCGAGGAACGGCGCAGTCACT 1955
Qy	421	TGCGCAGATGGCGCTTGCAATGGGTGCTGCGCGAG	CAAGAGGAGTACGGCGCGATACCGT 480
Db	1954	TGCGCAGATGGCGCTTGCAATGGGTGCTGCGCGAG	CAAGAGGAGTACGGCGCGATACCGT 1895
Qy	481	GACCAAGTCATTAATGTTGGTGTTCGTCAGTTAGAC	GTGGAACAACGCTTGATTCAT 540
Db	1894	GACCAAGTCATTAATGTTGGTGTTCGTCAGTTAGAC	GTGGAACAACGCTTGATTCAT 1835

Qy	541	CAACAACTTGGAGTTTCTGACGCGAGTTTGGAGCGGATCGATGAGATTTCCACGAGCG	600
Ds	1834	CAACAACTTGGAGTTTCTGACGCGAGTTTGGAGCGGATCGATGAGATTTCCACGAGCG	1775
Qy	601	CGGCATCAACATTTGGGCGAAGGCGACCGATTTCCAAACCCGCGAAACTAACCCATCAA	660
Ds	1774	CGGCATCAACATTTGGGCGAAGGCGACCGATTTCCAAACCCGCGAAACTAACCCATCAA	1715
Qy	661	CATCAGTTTGAATGCGCATGCGTTCATCACTGCGCACGACGAGTGTGATCGAGCGCA	720
Ds	1714	CATCAGTTTGAATGCGCATGCGTTCATCACTGCGCACGACGAGTGTGATCGAGCGCA	1655
Qy	721	CACCTTGGGCGTGGACGCGGCGTGACAAATGCTGCGCGCGAAACCCACACGCGGAA	780
Ds	1654	CACCTTGGGCGTGGACGCGGCGTGACAAATGCTGCGCGCGAAACCCACACGCGGAA	1595
Qy	781	CCAGATCAGGCTTGGCGGAAACGCGCGACGCGGCGAAATCCACGCTGCGGTGCGCGTA	840
Ds	1594	CCAGATCAGGCTTGGCGGAAACGCGCGACGCGGCGAAATCCACGCTGCGGTGCGCGTA	1535
Qy	841	TTGGCGCGCGACGCGCGCGATTAACACACAAACGCGTCCAAATACGCAATTCGGGTTCAACCA	900
Ds	1534	TTGGCGCGCGACGCGCGCGATTAACACACAAACGCGTCCAAATACGCAATTCGGGTTCAACCA	1475
Qy	901	GCTCAGCAGCATTTGCCATCAACATGCGCTTTTACCCAAACCGCGTCTGCTATCGAGCTCAC	960
Ds	1474	GCTCAGCAGCATTTGCCATCAACATGCGCTTTTACCCAAACCGCGTCTGCTATCGAGCTCAC	1415
Qy	961	CTCCACCGCGACCGGTTGCGGTCAGTGCGCACCGCGCGAAACCGCGCGGTC	1020
Ds	1414	CTCCACCGCGACCGGTTGCGGTCAGTGCGCACCGCGCGAAACCGCGCGGTC	1355
Qy	1021	ATCGGCGACGTTGCTGCTTCTCAATGATCTGCGGCTTCCACCTGTTGTCAT	1080
Ds	1354	ATCGGCGACGTTGCTGCTTCTCAATGATCTGCGGCTTCCACCTGTTGTCAT	1295
Qy	1081	GCGCTCTTTCGCTGCGCATGACGCGCAACCAATAACAGGTAAAGCGATGCCACCCAGCGCAT	1140
Ds	1294	GCGCTCTTTCGCTGCGCATGACGCGCAACCAATAACAGGTAAAGCGATGCCACCCAGCGCAT	1235
Qy	1141	AATATCAGACGATCGCGCGCGATTTGGAACAAAGATCAACCGCCCAAGGTGCGCGGAT	1200
Ds	1234	AATATCAGACGATCGCGCGCGATTTGGAACAAAGATCAACCGCCCAAGGTGCGCGGAT	1175
Qy	1201	GACCAAAAGACGTCAGAAATTAACACACAGAGAGAACCGCAATGAGTCTTCGCGCTT	1260
Ds	1174	GACCAAAAGACGTCAGAAATTAACACACAGAGAGAACCGCAATGAGTCTTCGCGCTT	1115
Qy	1261	AATTCCTGTTAATCACCAGTACATTTCTGCGTCCGATGGAACAGTAAAGACTGGCCCC	1320
Ds	1114	AATTCCTGTTAATCACCAGTACATTTCTGCGTCCGATGGAACAGTAAAGACTGGCCCC	1055
Qy	1321	CAAAAGCAGACTGTAATGAAGATTTCCATGATCACCATCGTACCTATGAAGTACTTA	1380
Ds	1054	CAAAAGCAGACTGTAATGAAGATTTCCATGATCACCATCGTACCTATGAAGTACTTA	995
Qy	1381	AGTAAATGATTGTTTCTTAACATGTTTAAATATAGTCTTCAACCCCATTTCACTGGA	1440
Ds	994	AGTAAATGATTGTTTCTTAACATGTTTAAATATAGTCTTCAACCCCATTTCACTGGA	935
Qy	1441	CACCTTCTCTCAATCATGATGAAGCAGCTTGAAGCGCCCTCTTAGCCCTTTCCAT	1500
Ds	934	CACCTTCTCTCAATCATGATGAAGCAGCTTGAAGCGCCCTCTTAGCCCTTTCCAT	875
Qy	1501	TTCCCTCTCGCGGTGAGTCAACGCTTAAAGCTCTGAGGATCAGTGGGTGAGTGT	1560
Ds	874	TTCCCTCTCGCGGTGAGTCAACGCTTAAAGCTCTGAGGATCAGTGGGTGAGTGT	815
Qy	1561	GGTATCGCGACCAACCGCGCAAGCAACGAGCGGTTGAAGTCTTTGTGCAAGCAGC	1620
Ds	814	GGTATCGCGACCAACCGCGCAAGCAACGAGCGGTTGAAGTCTTTGTGCAAGCAGC	755
Qy	1621	GGGAAAATGCTGCTGCTGAAGCAGAACTAAAGCGCAACTATCTGACGCGCTTGCTGA	1680
Ds	754	GGGAAAATGCTGCTGCTGAAGCAGAACTAAAGCGCAACTATCTGACGCGCTTGCTGA	695
Qy	1681	AATCCCGTTAAACATCGGCATCAACGAGATTCGCTATCAACATGTTTCTCCCGTGT	1740
Ds	694	AATCCCGTTAAACATCGGCATCAACGAGATTCGCTATCAACATGTTTCTCCCGTGT	635
Qy	1741	CAACGAGTGTGCTTCTGGGTTGGAGCAACGCTCAACGCTGCGCTTGAAGATGAAGCGCA	1800
Ds	634	CAACGAGTGTGCTTCTGGGTTGGAGCAACGCTCAACGCTGCGCTTGAAGATGAAGCGCA	575
Qy	1801	CACATTTATCTCTGCGCGCTGAGATGTTTGAAGAGCGTAAACCCGTTGAAGCTTAATCC	1860
Ds	574	CACATTTATCTCTGCGCGCTGAGATGTTTGAAGAGCGTAAACCCGTTGAAGCTTAATCC	515
Qy	1861	CGTGGCGGATGTGAAGTAGTAGAACTTGGAAACCATGCGGCACTTGGCCATTCGACCCC	1920
Ds	514	CGTGGCGGATGTGAAGTAGTAGAACTTGGAAACCATGCGGCACTTGGCCATTCGACCCC	455
Qy	1921	CTCATTTGCGGATGCTCATGTTGATGSGAACTAGATTGSGCTGCGATGCCGCTT	1980
Ds	454	CTCATTTGCGGATGCTCATGTTGATGSGAACTAGATTGSGCTGCGATGCCGCTT	395
Qy	1981	ACGCTTCGCTCCCAAGATGTCTTCAAGACCGTGAACCTGCAACGCGCGCTCGATGCTC	2040
Ds	394	ACGCTTCGCTCCCAAGATGTCTTCAAGACCGTGAACCTGCAACGCGCGCTCGATGCTC	335
Qy	2041	TGTGGGCGCAGCGCGGTATCCATTTGCTCGGTCGCGGAGGTGTTGTTGAGGCAATTCG	2100
Ds	334	TGTGGGCGCAGCGCGGTATCCATTTGCTCGGTCGCGGAGGTGTTGTTGAGGCAATTCG	275
Qy	2101	CCGAGGCTTGGTTGGGCACTTCTTCCGAAACCCCAAGCTGCTCCCATCTTAAAGCAGG	2160
Ds	274	CCGAGGCTTGGTTGGGCACTTCTTCCGAAACCCCAAGCTGCTCCCATCTTAAAGCAGG	215
Qy	2161	AGAGTGTATCTCTCGATGATGATACCATTCACACCGATGATTTGGCAACGATGGCG	2220
Ds	214	AGAGTGTATCTCTCGATGATGATACCATTCACACCGATGATTTGGCAACGATGGCG	155
Qy	2221	CCTGGATCTAGATCTCTAGCTAGATCTCACAGCGCGTGTGATGAGCAAGTGGG	2280
Ds	154	CCTGGATCTAGATCTCTAGCTAGATCTCACAGCGCGTGTGATGAGCAAGTGGG	95
Qy	2281	ATTGCGGCTTGTACTTCTGAAAAGGTTTCAAGGTTTTCATCTTCTGCGCGCAGGAA	2340
Ds	94	ATTGCGGCTTGTACTTCTGAAAAGGTTTCAAGGTTTTCATCTTCTGCGCGCAGGAA	35
Qy	2341	TTGGGCGCAGCAGATTAACCTTTCAGCAATGG	2374
Ds	34	TTGGGCGCAGCAGATTAACCTTTCAGCAATGG	1
RESULT 3			
AP005277			
LOCUS			
DEFINITION			
Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 4/10.			
ACCESSION			
AP005277 BA000036			
VERSION			
AP005277.1 GI:21323710			
KEYWORDS			
SOURCE			
ORGANISM			
Corynebacterium glutamicum ATCC 13032			
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
Corynebacterineae; Corynebacteriaceae; Corynebacterium.			
REFERENCE			
1 Nakagawa, S.			
Complete genomic sequence of Corynebacterium glutamicum ATCC 13032			
Unpublished			
2 (bases 1 to 333150)			
Nakagawa, S.			
Direct Submission			
Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.			

Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,  
Tokyo 194-8533, Japan (E-mail: enakagawa@axanagen.com,  
Tel: 81-44-829-3031, Fax: 81-44-813-1651)  
This sequence is conducted by collaboration of Kyowa Hakko Kogyo  
Co. Ltd. And Kitasato University.

## FEATURES Location/Qualifiers

```

source
  1..333150
    /organism="Corynebacterium glutamicum ATCC 13032"
    /mol_type="genomic DNA"
    /strain="ATCC 13032"
    /db_xref="taxon:196627"
    /note="ATCC 13032"
    /complement(63..1292)
    /gene="Cgl0944"
    /complement(63..1292)
    /gene="Cgl0944"
    /codon_start=1
    /transl_table=11
    /product="Hypothetical membrane protein"
    /protein_id="BAB98337.1"
    /db_xref="GI:21323711"
    /translation="MTNPTERRNARELIWANGLQNGDQIVAAKTVLPWLLQAAGAP
    FLALLVPIRRAGMLPQAAITGVLRQTSKVVWIGSGQFVSALGIGVAAALPLRG
    WALGITVIVLAAALSFRSMCSIAKQVQKVKISKRGGLVTRATVIGVGMGLVAGL
    AIALFGSHSPTRVLAADVAVTASSFSLFASIVFARIEYAKPATPKNANPWRRCI
    AAKDDKAFREIVVRSMMLVTALSTAFIVLAASGNSIDSLGFLFSLASGLASWVG
    RISGISWDSHSSKNVMAGGALFGSIVLILVLSAPAPQINTLVFPLFLITIAHTA
    IRVARTYVMDAGSDQRTYVADANTLMGVLLVGLSGLSGLFAGNEALLFLAAI
    GLIGTISARGLKEVSAG"
    1382..2317
    /gene="Cgl0945"
    1382..2317
    /gene="Cgl0945"
    /codon_start=1
    /transl_table=11
    /product="Putative multicopper oxidases"
    /protein_id="BAB98338.1"
    /db_xref="GI:21323712"
    /translation="MPELNRTFFKAGVLAATVVGAVLVACSSDDVRVGYGEPRTL
    PIPADLTREGSSVFALEAQGTGSOILPQVTKWGFNGTHLQFTLVVKKGDVHVH
    DVNNIDMTVHHGKMLPAIADGPHSPGQGTWSTVTVANDAAIWLHPHCHG
    LTGLHVRGLAGMLIIVDEATDKLDPREYGVDDIPLVMDHRFLEGLSDLELDPL
    GLGDPTTANGITWAFDATERVPRVLNGSNRNFYLAFSDTPTFQVLAISGSLD
    EQQRITLAIGFERWEIVLEPEGVDSVLESDYGVDEPDEDFEFGMSDSFLQ
    LITIGSDSAQALPGVLVKSTEDVDATEFTFMFTFSINDLQMDQWVDVVID
    HDQPEWIVTNDSDWPHNVHDARFKVLKEGTDELFDNDGMDTVGLPGGATAIL
    AVRFHYDPQWPYMYTHCHMLYHEDQGMGQFVIVREGDEPAVLGSGTSSIDSAGG
    HAH"
    3063..3983
    /gene="Cgl0946"
    3063..3983
    /gene="Cgl0946"
    /notes="PF00005:ABC transporter
    TIGR00960:3a0501a02: Type II (General) Secretory Pathway
    (ITSP) Family protein
    TIGR00968:3a0106801: sulfate transport system permease
    protein
    TIGR01166:cbiO: cobalt transport protein ATP-binding
    subunit
    TIGR01184:ntrCD: nitrate transport ATP-binding subunits C
    and D
    TIGR01186:prov: glycine betaine/L-proline transport ATP
    binding subunit
    TIGR01187:pota: spermidine/putrescine ABC transporter
    ATP-binding subunit
    TIGR01188:draA: daunorubicin resistance ABC transporter
    ATP-binding subunit
    TIGR01189:cmcA: heme exporter protein CmcA"
    /codon_start=1
    /transl_table=11
    /product="ABC-type transporter, ATPase component"
    /protein_id="BAB98339.1"
    /db_xref="GI:21323713"

```

```

    /translation="MTTTPAIDVTLVRTYGDTAVKGLNFHVQGEVFGLLGTNGAG
    KTSLEIVRGLSAGSGTIVISGLDEPVDRAILRPELGIIMLQSGGLPQSLVATMDM
    WHGCTVPAIKVDLADVLLHRENVKVGALSGQRRLDIALCIGPSPILFIDEPT
    GLDPSRRHTWQLLLDKORGVTWMLTTHYLEAEFLCDRIAIMNAGEIAVEGTILDE
    LVAREKSIISFVLRGGQVELPVLSSAEIIRONNHVRIATTTLQOHTLEILWAAETGI
    ALGFAAKPATLESVFMDIASLENTSLQTA"
    4003..4836
    /gene="Cgl0947"
    4003..4836
    /gene="Cgl0947"
    /codon_start=1
    /transl_table=11
    /product="ABC-type transporter, permease components"
    /protein_id="BAB98340.1"
    /db_xref="GI:21323714"
    /translation="MTTSHTARGLOHATPERKTSFFKTSLFKAELQFRNKTLFM
    ATPVPGIPLLLFLIGGSAESANSFDYFVMTLLVQFVTVLSMATTRDERVLAKR
    LRTEARDIIGALCFPGALLTLITVVIPLLMVWGAPAPINLVPVFAVLGLILL
    CSALMTSGFTRNAAQAQSMFVPLAMGGLSIRFPFGDSIVADILATYTPFAALS
    DLVQIGWAGATTADSVGGVEAANFAGIFQDMLIPLGILAAATAAVWAANRYMRWDSY
    R"
    4879..6138
    /gene="Cgl0948"
    4879..6138
    /gene="Cgl0948"
    /EC_number="2.7.3.-"
    /codon_start=1
    /transl_table=11
    /product="Two-component system, sensory transduction
    histidine kinases"
    /protein_id="BAB98341.1"
    /db_xref="GI:21323715"
    /translation="WNKDFMTAGTARWFSRGVSLASPVTAPLNSWRRLPNLAKYTL
    YTRVSIQAIPLVLLSAYFLGIVANAGTLNPSFWMLGFSVILLITVLVYQPSLNS
    HPRESVQPFETGLVNLVGVVSVVLOIPGLNSDMTRATALIFTLTCVLLGIAYI
    PMNRYWMLLIAMSAVLMTSTTVDLSALWVPIPLMAGTVLRSVTVDMKVEERS
    RELEASURVTEERLPAQELHDTLQHLAAMSVKSELALAKAGDRLNELRELQK
    LTRTSMSEMDVVGRTVNLATEIEGAKSLADAHHLHSVITGTVTSQVSPHARELCAM
    LVREATTNLRHSDATATLTLSSTEVMDNNGVNDKIDRLSGLSALSRAESAGMTL
    IVSREDDQFSVRMLINAPANTPAEKEA"
    6139..6750
    /gene="Cgl0949"
    6139..6750
    /gene="Cgl0949"
    /notes="PF00072:Response regulator receiver domain
    PF00196:Bacterial regulatory proteins, luxR family"
    /codon_start=1
    /transl_table=11
    /product="Two-component system, response regulators
    consisting of a CheY-like receiver domain and a HTH
    DNA-binding domain"
    /protein_id="BAB98342.1"
    /db_xref="GI:21323716"
    /translation="MISITADDEALTIASSLATLISLEPDLVPTAGSGEELIETWA
    DPSNRDVCVLQGLGIDGIDTATRLMETPDIAVLIVTSHPROLKRALAAGVLG
    ZLPKTSDEPATARVTHAGRYIIDPELAAMTISAGESPLTNRESEVLELAGQLSA
    EETAAHLAPGTTNYLSQMTKVGQNRFEATFRAREGLML"
    complement(6781..6987)
    /gene="Cgl0950"
    complement(6781..6987)
    /gene="Cgl0950"
    /codon_start=1
    /transl_table=11
    /product="Hypothetical protein"
    /protein_id="BAB98343.1"
    /db_xref="GI:21323717"
    /translation="MKHWSMYGAGPLNTPHTKVEVFMDSIVNVHFWFSTNNWVN
    MTGLPRETFVPAFVGSITVLVLSF"
    7048..8205
    /gene="Cgl0951"
    7048..8205
    /gene="Cgl0951"
    /codon_start=1

```

/transl_table=11	
/product="tRNA-processing ribonuclease BN"	
/protein_id="BAB98344.1"	
/db_xref="GI:21232718"	
/translation="MTAMUCSFPSGVQDEYDMNGVQVQEHLDATLIIADHFNPN	
SDGRKERINFQWKYALNTRVDVDFPDGLDAAALTFSTLSAPAVILGYSVITIP	
LASDSTELANLVDENVQYVPEQSHVNGVIDSIAGSAAAGQVGVAVGVIITALMTSS	
AVRASFCANAVYGRSEGRITIKRWAMLLFNLLALLGIIILVSNVLTNEVLMGIP	
APIASPLHNTVLSFLTDREPMIIMWRFPIVIVGLIMFVATLYWAPNARPKFRML	
SLGSFLAIVGILLAGVLFNFTLPAFAFSYSGVAGSLIAVFLAIFWPNICLIIGIKID	
VEISRAQQLQMGAGSLVLPFRSIEKVAQKQKQRLMDQAAARESN"	
8672..10114	
gene	/gene="Cg10952"
8672..10114	
CDS	/gene="Cg10952"
/note="PF00005:ABC transporter	
pf00664:ABC transporter transmembrane region."	
/codon_start=1	
/transl_table=11	
/product="ABC-type multidrug/protein/lipid transport	
system, ATPase component"	
/protein_id="BAB98345.1"	
/db_xref="GI:21232719"	
/translation="MMRTHTASSILRRMTRRQKVAFGAFFLGMQMLSEALVPIAI	
GLIVDHALTKDLRLVGLVAFVVLVFLVSFYRFGSRALNRVNPESHARVEVAD	
HALKNDLPRNLVPEVMSSTADSDSTRIFQIGITGVSAATGFLGAATYLLISDLV	
GLILVLPITISGVVALASKGISKRSVTOEKLAESGAQSDIMWGLRVITKAIGERW	
AVKTEKASQASARAADTVASGKVAIGELSLIVNLAAVLLAGWRVITGELPGGO	
LIAIVGVASTLYSEPIRLSNSINASATAGHAARVAFNLDESQAQYSSSTINDGE	
FLIVVAVPTLPHGNDILATPHADIFEGTIRSNISMHEDNVPIPOVIRASGUTDI	
IIVGDLAPVDTVSSGSLGQKQVALLARALHADAEVLVMDPTSAVDVSVEYSIAQG	
IKQLRAGKTTIVSSSPAFYNLADRVISHV"	
10107..11843	
gene	/gene="Cg10953"
10107..11843	
CDS	/gene="Cg10953"
Query Match	
Best Local Similarity 100.0%; Score 2374; DB 1; Length 333150;	
Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 AGATACCTCTTTGGAAGAACCATGTACCATTCGGTGACATTTGGTGTGCTGGAAGGC 60
Db	327645 AGATACCTCTTTGGAAGAACCATGTACCATTCGGTGACATTTGGTGTGCTGGAAGGC 327704
Qy	61 TCTTTACGTGGGTATTTCTTCTACGGTCCAGAGCTCACAGCGGAGCGGCGGTGAGTTCAAT 120
Db	327705 TCTTTACGTGGGTATTTCTTCTACGGTCCAGAGCTCACAGCGGAGCGGCGGTGAGTTCAAT 327764
Qy	121 GCGGAGGAGGCGTCCGCGTTCTGATTATCATCAGCAAGCTATTCATCATTAATCGTTG 180
Db	327765 GCGGAGGAGGCGTCCGCGTTCTGATTATCATCAGCAAGCTATTCATCATTAATCGTTG 327824
Qy	181 GGTGGAGAACCGGGCGATGACGTGAGAACTTGTTCAGTCAAGTCCAGTCCCAATGTCCT 240
Db	327825 GGTGGAGAACCGGGCGATGACGTGAGAACTTGTTCAGTCAAGTCCAGTCCCAATGTCCT 327884
Qy	241 TGGCGTCAATTTGCTTTTCTCACCACTTGGCAGGCGCTGTCAACGGAATAATCTCGATGG 300
Db	327885 TGGCGTCAATTTGCTTTTCTCACCACTTGGCAGGCGCTGTCAACGGAATAATCTCGATGG 327944
Qy	301 AATTCCAGAGGGTTCGCGCCAGCAGCGGTGAAGTCCCTGTCTGAGGGCATGTTGAACGT 360
Db	327945 AATTCCAGAGGGTTCGCGCCAGCAGCGGTGAAGTCCCTGTCTGAGGGCATGTTGAACGT 328004
Qy	361 GAACAATATTGATGGTCCGACAGCTCAATGACATGCCACAGGAACGCGGAGTCACT 420
Db	328005 GAACAATATTGATGGTCCGACAGCTCAATGACATGCCACAGGAACGCGGAGTCACT 328064
Qy	421 TCGCAGATCGGCGCTTGCATGGGTGCTGCGGAGCAAGAGAGTACGCGCGGATACCGT 480
Db	328065 TCGCAGATCGGCGCTTGCATGGGTGCTGCGGAGCAAGAGAGTACGCGCGGATACCGT 328124

Qy	481 GACCAGTGCATTGATTGGTGTCTCGTCACTTGAGCAGCTGGACAAACAGCGTTGATTCACT 540
Db	328125 GACCAGTGCATTGATTGGTGTCTCGTCACTTGAGCAGCTGGACAAACAGCGTTGATTCACT 328184
Qy	541 CAACAACTTGGAGTGTCTGACCGCGAGTTGGAGCGCATCGATGAGATTTCCCAAGCGC 600
Db	328185 CAACAACTTGGAGTGTCTGACCGCGAGTTGGAGCGCATCGATGAGATTTCCCAAGCGC 328244
Qy	601 CGGCATCAACATTTGGGGAAGGCCACCGATTCCAAAACCCCGGAAACTAACCCATCAA 660
Db	328245 CGGCATCAACATTTGGGGAAGGCCACCGATTCCAAAACCCCGGAAACTAACCCATCAA 328304
Qy	661 CATCAGTTTGTATGGCCAAATCGCGTCAATCAAACTGCGCAACGACGCTTGATCCAGCGCCA 720
Db	328305 CATCAGTTTGTATGGCCAAATCGCGTCAATCAAACTGCGCAACGACGCTTGATCCAGCGCCA 328364
Qy	721 CACCTTGGGCGTGGACACGCGGGGTGACAAATGCTGCTGGCGGAAACCCACCGCGGAA 780
Db	328365 CACCTTGGGCGTGGACACGCGGGGTGACAAATGCTGCTGGCGGAAACCCACCGCGGAA 328424
Qy	781 CCAGATCAGGCTTTGGCGGAAACCGCGCAGCGGCAAAAATCCACCGTCCGGTGTGCGCGTA 840
Db	328425 CCAGATCAGGCTTTGGCGGAAACCGCGCAGCGGCAAAAATCCACCGTCCGGTGTGCGCGTA 328484
Qy	841 TTGCGCGCGACGCCCGCGCATAAACAAACCGGTCCAAATAGCGATTCGGGTTCAACCA 900
Db	328485 TTGCGCGCGCGACGCCCGCGCATAAACAAACCGGTCCAAATAGCGATTCGGGTTCAACCA 328544
Qy	901 GGTGAGCAGCATGTCATCAACATGGGCTTTACCCAAACCGGTGCTTATCGAGCTCAC 960
Db	328545 GGTGAGCAGCATGTCATCAACATGGGCTTTACCCAAACCGGTGCTTATCGAGCTCAC 328604
Qy	961 CTCCACCGCGCACCCGGTTCGCGGTGTGTCAGTGGCCACCGCGGAAACCCCAAGGCGTGC 1020
Db	328605 CTCCACCGCGCACCCGGTTCGCGGTGTGTCAGTGGCCACCGCGGAAACCCCAAGGCGTGC 328664
Qy	1021 ATCGGCGCACCGTTGGTTCGTTTCTCAATGATCTGTGGGCGCTTCCACCTTGTGTCAT 1080
Db	328665 ATCGGCGCACCGTTGGTTCGTTTCTCAATGATCTGTGGGCGCTTCCACCTTGTGTCAT 328724
Qy	1081 GCGCTCTTTCGCTGCGCATGACGCAACCATTAACAGGTAAAGCATGCCACCCAGCGCAT 1140
Db	328725 GCGCTCTTTCGCTGCGCATGACGCAACCATTAACAGGTAAAGCATGCCACCCAGCGCAT 328784
Qy	1141 AATATCGAGCAAGATCGGCGCGCATTTGGACAAAAGATCAACGCGCAAGGTGCGGCGAT 1200
Db	328785 AATATCGAGCAAGATCGGCGCGCATTTGGACAAAAGATCAACGCGCAAGGTGCGGCGAT 328844
Qy	1201 GAACAAAAGAGCTGAGAAATTAACACACGAGAGAACCGCAATGAGTCTTTCGCGCTT 1260
Db	328845 GAACAAAAGAGCTGAGAAATTAACACACGAGAGAACCGCAATGAGTCTTTCGCGCTT 328904
Qy	1261 AATTCTTGTGTTAATCAACAGTACATTTCTGCGGTCCGATGGACAGTAAAGAGTGGCCCC 1320
Db	328905 AATTCTTGTGTTAATCAACAGTACATTTCTGCGGTCCGATGGACAGTAAAGAGTGGCCCC 328964
Qy	1321 CAAAAGCAGACCTGTGTAATGAGAGATTCATGATACCATGTGACCTATGGAAGTACTTA 1380
Db	328965 CAAAAGCAGACCTGTGTAATGAGAGATTCATGATACCATGTGACCTATGGAAGTACTTA 329024
Qy	1381 AGTAAATGATGTTCTTAAATGAGTAAATATAGTTCATGAACCCATTCAACTGGA 1440
Db	329025 AGTAAATGATGTTCTTAAATGAGTAAATATAGTTCATGAACCCATTCAACTGGA 329084
Qy	1441 CACTTTGCTCTCAATCATTTGATGAAGCAGCTTCGAAGGCGCTTCCTTAGCGCTTTCAT 1500
Db	329085 CACTTTGCTCTCAATCATTTGATGAAGCAGCTTCGAAGGCGCTTCCTTAGCGCTTTCAT 329144
Qy	1501 TTCCTCCCTCGCGGTGAGTCAAGCGGTTAAAGCTTCGAGCATCAGTGGTTCGAGTGT 1560
Db	329145 TTCCTCCCTCGCGGTGAGTCAAGCGGTTAAAGCTTCGAGCATCAGTGGTTCGAGTGT 329204
Qy	1561 GGTATCGCGCACCCCAACCGGCGCAAGCAACCGAGCGGTGAGTCTTGTGCAAGCAGC 1620



[illegible]



1381 ACTAATGATGTTGGTCTTACATGGTTTAAATATAGTCTCATGAACCCGATTTCAACTGGA 1440  
Db |||||  
61 AGTAAATGATGTTGGTCTTAAATGTTTAAATATAGTCTCATGAACCCGATTTCAACTGGA 120  
Qy |||||  
1441 CACTTTGCTCTCAATCATGATGAAGCAGCTTGAAGCGGCTCTTACGCTTTCAT 1500  
Db |||||  
121 CACTTTGCTCTCAATCATGATGAAGCAGCTTGAAGCGGCTCTTACGCTTTCAT 180  
Qy |||||  
1501 TTCCCGCTCGCGGCTGAGTCAGCGCTTAAAGCTCTCAGCATCACTGGTTCGAGTGT 1560  
Db |||||  
181 TTCCCGCTCGCGGCTGAGTCAGCGCTTAAAGCTCTCAGCATCACTGGTTCGAGTGT 240  
Qy |||||  
1561 GGTATCGCGCACCACCGCGGCTGAAGCAACCGAGCGGCTGAAGTCTTGTGCAAGCAGC 1620  
Db |||||  
241 GGTATCGCGCACCACCGCGGCTGAAGCAACCGAGCGGCTGAAGTCTTGTGCAAGCAGC 300  
Qy |||||  
1621 GCGGAAATGCTGCTGCAAGCAGAACTAAAGCGCAACTATCTGAGCGCTTGTCTGA 1680  
Db |||||  
301 GCGGAAATGCTGCTGCAAGCAGAACTAAAGCGCAACTATCTGAGCGCTTGTCTGA 360  
Qy |||||  
1681 AATCCCGTTAAACCATGCCATCAACGAGATTCGCTATCCATGCTTTCCTCCGCTGT 1740  
Db |||||  
361 AATCCCGTTAAACCATGCCATCAACGAGATTCGCTATCCATGCTTTCCTCCGCTGT 420  
Qy |||||  
1741 CAACGAGTAGCTTCTTGGGGTGGAGCAACGCTCAAGCTGCGCTTGAAGTGAAGCGCA 1800  
Db |||||  
421 CAACGAGTAGCTTCTTGGGGTGGAGCAACGCTCAAGCTGCGCTTGAAGTGAAGCGCA 480  
Qy |||||  
1801 CACATATCTCTGCGCGGCTGGAGATGTTTTAGAGCGGTAAACCGTGAAGCTTAATCC 1860  
Db |||||  
481 CACATATCTCTGCGCGGCTGGAGATGTTTTAGAGCGGTAAACCGTGAAGCTTAATCC 540  
Qy |||||  
1861 CGTGGCGGATGTAAGTAGTAGTAAGCTTGGAAACGATGCGGCTTGGCAATTCGAACCCC 1920  
Db |||||  
541 CGTGGCGGATGTAAGTAGTAGTAAGCTTGGAAACGATGCGGCTTGGCAATTCGAACCCC 600  
Qy |||||  
1921 CTCAATGCGGATGCTTACATGTTGATGGAACTAGATTCGGCTGCGATGCGCTTCT 1980  
Db |||||  
601 CTCAATGCGGATGCTTACATGTTGATGGAACTAGATTCGGCTGCGATGCGCTTCT 660  
Qy |||||  
1981 ACCTTTCGCTCCCAAGATGCTTCAAGACCGTGAACCTGACCGGCTGATGCTTCC 2040  
Db |||||  
661 ACCTTTCGCTCCCAAGATGCTTCAAGACCGTGAACCTGACCGGCTGATGCTTCC 720  
Qy |||||  
2041 TGTGGGCGCAGCGCTATCCATGTTCCGCTGCGGAGGTTTGTGAGCAATTCG 2100  
Db |||||  
721 TGTGGGCGCAGCGCTATCCATGTTCCGCTGCGGAGGTTTGTGAGCAATTCG 780  
Qy |||||  
2101 CCGAGGCTTGGTGGGACTTCTCCGAAACCAAGCTGCTCCCATGCTTAAAGCAGG 2160  
Db |||||  
781 CCGAGGCTTGGTGGGACTTCTCCGAAACCAAGCTGCTCCCATGCTTAAAGCAGG 840  
Qy |||||  
2161 AGAAGTATCTCTTCGATGAGATACCAATGACACCGATGATTTGGCAACGATGGCG 2220  
Db |||||  
841 AGAAGTATCTCTTCGATGAGATACCAATGACACCGATGATTTGGCAACGATGGCG 900  
Qy |||||  
2221 CCGAATCTAGATCTTCTAGCTAGCTACACGACCGCTGTTGATGAGCAATTCGAGG 2280  
Db |||||  
901 CCGAATCTAGATCTTCTAGCTAGCTACACGACCGCTGTTGATGAGCAATTCGAGG 960  
Qy |||||  
2281 ATTGGCGCTTACTTCTTCTGAAAGGTTTCAAG 2313  
Db |||||  
961 ATTGGCGCTTACTTCTTCTGAAAGGTTTCAAG 993

RESULT 6  
AX244055  
LOCUS AX244055 993 bp DNA linear PAT 29-SEP-2001  
DEFINITION Sequence 47 from Patent WO0166573.  
ACCESSION AX244055  
VERSION AX244055.1 GI:15959119  
KEYWORDS  
SOURCE Corynebacterium glutamicum

Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
1  
Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O., Haberhauer, G.,  
Kim, J.W., Lee, H.S. and Hwang, B.J. 2001.  
Corynebacterium glutamicum genes encoding metabolic pathway  
proteins  
Patent: WO 0166573-A 47 13-SEP-2001;  
BASF AKTIENGESSELLSCHAFT (DE)  
Location/Qualifiers  
1. .993  
/organism="Corynebacterium glutamicum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1718"  
101. .973  
/note="unnamed protein product; RXA01393"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC88506.1"  
/db\_xref="GI:15959120"  
/translation="NNPQLDTLISIIDEGSPFGASLALSISSAVSORVKALEHHVG  
RVLSRTPAKATEAGVLAQARKVLLQAEKQSLGRLAIPLTIAINADLSLSTW  
FPPVNEVAVSGAGTILRLDEAHTLSLRGDLVGAVTREANPVAGCEVVELTWR  
HLATLSLRDADTMDVKGKLDWAAMPVLRFGPKDVLQEDLDGVDGSPVGRREVSVPS  
AEGFGEAIRREGLGWGLLPETQAAPMLKAGEVILLDEIPIDTPHWWRKLESRLRL  
TDAVVDAAIEGLRP"  
Query Match 41.88; Score 993; DB 6; Length 993;  
Best Local Similarity 100.0%; Pred. No. 5.4e-233; Indels 0; Gaps 0;  
Matches 993; Conservative 0; Mismatches 0;  
1371 CAAAGCAGACCTGTAAATGAAGATTTCCATGATCACCATCGTGACCTATGAAGTACTTA 1380  
Db 1 CAAAGCAGACCTGTAAATGAAGATTTCCATGATCACCATCGTGACCTATGAAGTACTTA 60  
Qy 1381 AGTAAATGATGTTGTTCTTAAACATGATTAATAGCTTCTCAAGACCCCATTCACCTGA 1440  
Db 61 AGTAAATGATGTTGTTCTTAAACATGATTAATAGCTTCTCAAGACCCCATTCACCTGA 120  
Qy 1441 CACTTTGCTCTCAATCATGATGAAGCAGCTTGAAGCGGCTCTTACGCTTTCAT 1500  
Db 121 CACTTTGCTCTCAATCATGATGAAGCAGCTTGAAGCGGCTCTTACGCTTTCAT 180  
Qy 1501 TTCCCGCTCGCGGCTGAGTCAGCGCTTAAAGCTCTCAGCATCACTGGTTCGAGTGT 1560  
Db 181 TTCCCGCTCGCGGCTGAGTCAGCGCTTAAAGCTCTCAGCATCACTGGTTCGAGTGT 240  
Qy 1561 GGTATCGCGCACCACCGCGGCTGAAGCAACCGAGCGGCTGAAGTCTTGTGCAAGCAGC 1620  
Db 241 GGTATCGCGCACCACCGCGGCTGAAGCAACCGAGCGGCTGAAGTCTTGTGCAAGCAGC 300  
Qy 1621 GCGGAAATGCTGCTGCAAGCAGAACTAAAGCGCAACTATCTGAGCGCTTGTCTGA 1680  
Db 301 GCGGAAATGCTGCTGCAAGCAGAACTAAAGCGCAACTATCTGAGCGCTTGTCTGA 360  
Qy 1681 AATCCCGTTAAACCATGCCATCAACGAGATTCGCTATCCATGCTTTCCTCCGCTGT 1740  
Db 361 AATCCCGTTAAACCATGCCATCAACGAGATTCGCTATCCATGCTTTCCTCCGCTGT 420  
Qy 1741 CAAAGTAGTACTTCTTGGGCTGGAGCAACGCTCAAGCTGCGCTTGAAGTGAAGCGCA 1800  
Db 421 CAAAGTAGTACTTCTTGGGCTGGAGCAACGCTCAAGCTGCGCTTGAAGTGAAGCGCA 480  
Qy 1801 CACATATCTCTGCGCGGCTGGAGATGTTTTAGAGCGGTAAACCGTGAAGCTTAATCC 1860  
Db 481 CACATATCTCTGCGCGGCTGGAGATGTTTTAGAGCGGTAAACCGTGAAGCTTAATCC 540  
Qy 1861 CGTGGCGGATGTAAGTAGTAGTAAGCTTGGAAACGATGCGGCTTGGCAATTCGAACCCC 1920  
Db 541 CGTGGCGGATGTAAGTAGTAGTAAGCTTGGAAACGATGCGGCTTGGCAATTCGAACCCC 600

1921 CTCAATGGGGGATGCTTACATGTTGATPGGAAATAGATTTGGGCTCGATGCCGCTTT 1980  
 Db CTCAATGGGGGATGCTTACATGTTGATPGGAAATAGATTTGGGCTCGATGCCGCTTT 660  
 1981 AGCTTGGTCCAAAGATGCTTCAAGACCGTGAACCTGACGCGCGCTCGATGCTCC 2040  
 Db AGCTTGGTCCAAAGATGCTTCAAGACCGTGAACCTGACGCGCGCTCGATGCTCC 720  
 2041 TGTGGGCGCGAGCGCTATCCATTTGCCCTCGCGGGAAGTTTGTGAGCAATTCG 2100  
 Db TGTGGGCGCGAGCGCTATCCATTTGCCCTCGCGGGAAGTTTGTGAGCAATTCG 780  
 2101 CCGAGGCTTGGTTGGGGAATCTTCCCGAAACCAAGCTGCTCCCATGCTTAAAGCAG 2160  
 Db CCGAGGCTTGGTTGGGGAATCTTCCCGAAACCAAGCTGCTCCCATGCTTAAAGCAG 840  
 2161 AGAAGTATCTCTCCGATGAGTACCATTTGACACCGATGATTTGGCAACGATGGG 2220  
 Db AGAAGTATCTCTCCGATGAGTACCATTTGACACCGATGATTTGGCAACGATGGG 900  
 2221 CCTGGAATCTAGATCTTAGTACTCAGACGCGCTGTTGATGAGCAATTCGAGG 2280  
 Db CCTGGAATCTAGATCTTAGTACTCAGACGCGCTGTTGATGAGCAATTCGAGG 960  
 2281 ATTGGGCTTACTTACTTCTGAAAAGTTTCAG 2313  
 Db ATTGGGCTTACTTACTTCTGAAAAGTTTCAG 993

RESULT 7  
 AX123540 870 bp DNA linear PAT 11-MAY-2001  
 LOCUS Sequence 3456 from Patent EP1108790.  
 DEFINITION AX123540  
 ACCESSION AX123540  
 VERSION AX123540.1 GI:14041028  
 KEYWORDS Corynebacterium glutamicum  
 SOURCE Corynebacterium glutamicum  
 ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 1  
 REFERENCE Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,  
 AUTHORS Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.  
 TITLE Novel polynucleotides  
 JOURNAL Patent: EP 1108790-A 3456 20-JUN-2001.  
 KYOWA HAKKO KOGYO CO., LTD. (JP)  
 FEATURES Location/Qualifiers  
 source 1..870  
 /organism="Corynebacterium glutamicum"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1718"

ORIGIN  
 Query Match 36.6%; Score 870; DB 6; Length 870;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-203;  
 Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1421 ATGAACCCCAATCACTGACACTTTGCTCTCAATATTGATGAAGGAGCTTCGAGGC 1480  
 Db 1 ATGAACCCCAATCACTGACACTTTGCTCTCAATATTGATGAAGGAGCTTCGAGGC 60  
 1481 GCTTCCTTAGCCCTTCATTTCCCTCGGCGGTGAGTCAGCGGTTAAAGCTTCGAG 1540  
 Db 61 GCTTCCTTAGCCCTTCATTTCCCTCGGCGGTGAGTCAGCGGTTAAAGCTTCGAG 120  
 1541 CATCAGTGGGTGAGTGTGATATCGCCACCCAAACCGCCAAAGCAACCGAGCGGT 1600  
 Db 121 CATCAGTGGGTGAGTGTGATATCGCCACCCAAACCGCCAAAGCAACCGAGCGGT 180  
 1601 GAAGTCTTGTGCAAGCAGCGGAAATGTTGCTGCAAGCAAACTTAAGCGCA 1660  
 Db 181 GAAGTCTTGTGCAAGCAGCGGAAATGTTGCTGCAAGCAAACTTAAGCGCA 240

1661 CTATCTGACGCGCTTGTGTAATCCGTTAAACCATCGCATCAACGAGATTGCTATCC 1720  
 Db CTATCTGACGCGCTTGTGTAATCCGTTAAACCATCGCATCAACGAGATTGCTATCC 300  
 1721 ACATGTTTCTCCCGCTTCAACGAGGTAGCTTCTTGGGTGGAGCAAGCTCAGCGTG 1780  
 Db ACATGTTTCTCCCGCTTCAACGAGGTAGCTTCTTGGGTGGAGCAAGCTCAGCGTG 360  
 1781 CGCTTGGAGATGAAGCGCACATTTATCTTGTGCGGCTGGAGATGTTTGAAGCG 1840  
 Db CGCTTGGAGATGAAGCGCACATTTATCTTGTGCGGCTGGAGATGTTTGAAGCG 420  
 1841 GTAAACCGGTGAAGCTTAATCCGCTGGCGGATGTAAGTAGTAGTCTTGGAACTGG 1900  
 Db GTAAACCGGTGAAGCTTAATCCGCTGGCGGATGTAAGTAGTAGTCTTGGAACTGG 480  
 1901 CACTTGGCCATTTGCAACCCCTCATTTGCGGATGCTTACATGTTGATGGGAACTAG 1960  
 Db CACTTGGCCATTTGCAACCCCTCATTTGCGGATGCTTACATGTTGATGGGAACTAG 540  
 1961 TGGGCTCGGATGCGGCTTCTTACGCTTCCAAAGATGTTCTCAAGACGCTGAC 2020  
 Db TGGGCTCGGATGCGGCTTCTTACGCTTCCAAAGATGTTCTCAAGACGCTGAC 600  
 2021 GACGCGCGCTGATGCTTCTTGGGCGGCGGCTTATCCATTGTCCTGCGCGGAA 2080  
 Db GACGCGCGCTGATGCTTCTTGGGCGGCGGCTTATCCATTGTCCTGCGCGGAA 660  
 2081 GGTGTTGCTGAGCGCAATTCGCGGAGGCTTGTGTTGGGACCTTCTCCGAAACCC 2140  
 Db GGTGTTGCTGAGCGCAATTCGCGGAGGCTTGTGTTGGGACCTTCTCCGAAACCC 720  
 2141 GCTCCCATGCTTAAAGCAGGAGAGTATCTTCTCGATGATGATACCATTCACACCG 2200  
 Db GCTCCCATGCTTAAAGCAGGAGAGTATCTTCTCGATGATGATACCATTCACACCG 780  
 2201 ATGTTATGGCAAGTGGCGCTGGAATCTAGATCTCTAGCTACAGACGCGCTC 2260  
 Db ATGTTATGGCAAGTGGCGCTGGAATCTAGATCTCTAGCTACAGACGCGCTC 840  
 2261 GTTGATGACGCAATTCGAGGATTCGCGCT 2290  
 Db GTTGATGACGCAATTCGAGGATTCGCGCT 870

RESULT 8  
 BD165657 870 bp DNA linear PAT 17-JAN-2003  
 LOCUS BD165657  
 DEFINITION Novel polynucleotide.  
 ACCESSION BD165657  
 VERSION BD165657.1 GI:27871469  
 KEYWORDS JP 2002191370-A/3456.  
 SOURCE unidentified  
 ORGANISM unclassified.  
 1 (bases 1 to 870)  
 REFERENCE Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,  
 AUTHORS Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.  
 TITLE Novel polynucleotide  
 JOURNAL Patent: JP 2002191370-A 3456 09-JUL-2002;  
 KYOWA HAKKO KOGYO CO LTD  
 COMMENT OS Corynebacterium glutamicum  
 PN JP 2002191370-A/3456  
 PD 09-JUL-2002  
 PF 15-DEC-2000 JP 2000405096  
 PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI,  
 PI KEIKO OCHIAI,  
 PI HARUHIKO YOKOI,NAOKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO  
 PI OZAKI  
 PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC  
 C12N1/15,  
 PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/



Qy	2201	ATGTTATTGGCGAAAGATGGCGGCTCGGAATCTAGATCTTACGTAGACTCAACAGACGCCGTC	2265	
Db	781	ATGTTATTGGCGAAAGATGGCGGCTCGGAATCTAGATCTTACGTAGACTCAACAGACGCCGTC	840	
Qy	2261	GTGTATGCAGCAATCGAGGATTCGGGCGCT	2290	
Db	841	GTGTATGCAGCAATCGAGGATTCGGGCGCT	870	
RESULT 9	AX063771/c	822 bp	linear	PAT 24-JAN-2001
LOCUS	AX063771	Sequence 53 from Patent WO0100843.		
DEFINITION	AX063771			
ACCESSION	AX063771.1	GI:12541483		
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
CDS				
ORIGIN				
Query Match	34.6%	Score 822;	DB 6;	Length 822;
Best Local Similarity	100.08;	Pred. No. 6.1e-191;		
Matches 822;	Conservative	0;	Mismatches	0;
			Indels	0;
Gaps	0;			
Qy	629	GATTCACAAACCGGAAACCAACCCATCAACATCAGTTGATGGCCCAATGCGGTATC	698	
Db	822	GATTCACAAACCGGAAACCAACCCATCAACATCAGTTGATGGCCCAATGCGGTATC	763	
Qy	689	ACAACTGCCACGACGACGTTGATTCAGGCGCCACACCTTGGGCGCTGGACAGCGGGCGGTGAC	748	
Db	762	ACAACTGCCACGACGACGTTGATTCAGGCGCCACACCTTGGGCGCTGGACAGCGGGCGGTGAC	703	
Qy	749	AATGTGTCGCGCGGAAACCCACAGCAGCGGGAAACAGATCAGGCTTTCGCGGAAACCGCGCA	808	
Db	702	AATGTGTCGCGCGGAAACCCACAGCAGCGGGAAACAGATCAGGCTTTCGCGGAAACCGCGCA	643	
Qy	809	GGCGGAAATCCAGTCGCGGTGTCGCGGTATTGCGCGCGACGCGCGCGTAACACA	868	
Db	642	GGCGGAAATCCAGTCGCGGTGTCGCGGTATTGCGCGCGACGCGCGCGTAACACA	593	
Qy	869	AACGGCTCAAAATGACATTCGGGTTCACCAACGAGTCAGCAGATTCGATCAACATGGGC	928	
Db	582	AACGGCTCAAAATGACATTCGGGTTCACCAACGAGTCAGCAGATTCGATCAACATGGGC	523	
Qy	929	TTTACCCAAACCCGGTCTTATCGACGCTCACTTCCACCCGCAACCCGGTTGCGGTGCTCA	988	
Db	522	TTTACCCAAACCCGGTCTTATCGACGCTCACTTCCACCCGCAACCCGGTTGCGGTGCTCA	463	

[illegible]

Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and Gojobori, T.  
 Comparative Complete Genome Sequence Analysis of the Amino Acid Replacements Responsible for the Thermostability of *Corynebacterium efficiens*  
 Genome Res. 13 (7), 1572-1579 (2003)  
 22723752  
 12840036  
 2 (bases 1 to 308650)  
 Kawarabayashi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and Director-General of Biotechnology Center.  
 Direct Submission  
 Submitted (17-MAY-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan  
 (E-mail: bioente.go.jp, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)  
 Kawarabayashi, Y. is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, 305-8566 Japan  
 Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan  
 Itoh, T. is at the Japan Biological Information Research Center, Koto-ku, Tokyo, 135-0064 Japan  
 Yamagishi, A. is at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392 Japan  
 Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co., Inc., Kawasaki, Kanagawa, 210-8681 Japan  
 The other authors are at the National Institute of Technology and Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.  
 Location/Qualifiers  
 1..308650  
 /organism="Corynebacterium efficiens YS-314"  
 /mol\_type="genomic DNA"  
 /strain="YS-314"  
 /db\_xref="taxon:196164"  
 complement(83)..241)  
 /note="CE1178"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAC17989.1"  
 /db\_xref="GI:23493017"  
 /translation="MTTGIILGTGFGDAGLAKIGSDIPRGVGIHFRKPPVALQYR  
 TGAGARNQ"  
 365..742  
 /note="CE1179"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAC17989.1"  
 /db\_xref="GI:23493018"  
 /translation="WEKSQLOKQDFASVDHLSDDAAAFIDNLSRGAMHRAFLH  
 LVHACRQELARQDVFVTECRNDNVTPASDLKAKLSLANECAPGGAEDAAI  
 QRPESFAKVSQVRAVRRTQGR"  
 774..1271  
 /note="CE1180, similar to A0292256-1|CAC21724.1| percent  
 identity: 26 in 155 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="BAC17990.1"  
 /db\_xref="GI:23493019"  
 /translation="WFSNGVGGIILLFIVGLVIGPERLPRIMEDLKAALVAARTAI  
 NNAQNLDTFGEEDERKPMQIAQIRSMSPKTAITRALFGDSDSFLDSFDPKKIM  
 ADDTAGARQQVAKTGTTEHAEVVERPADRIDPTQSPNTFGKNDGKPPAGGYT  
 WTDLI"  
 complement(1353)..2480)  
 /note="CE1181, similar to AL390975-15|CAC01353.1| percent  
 identity: 63 in 370 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative ATP-binding protein Mrp"  
 /protein\_id="BAC17991.1"

/db\_xref="GI:23493020"  
 /translation="WMSQVTEASVRSALSRVEDPEIGKPIKTELGMKSVAGDSVQV  
 EYVITIAACPKMTIIVSTNTEAIIKIDIEGVGVQVITVMSDEQRRLRYSRGNTADP  
 VTPAQRGTRVYAVAGSGVKSMTNLTALAAKGLSVGLDADIYGHSPFGM  
 LSSAPRQHVDDIMPPQAGHGVKMSIATHFVGNAPVVRGPMHLRAIQOFLSDVSPFG  
 DDIILLDPPGTGDIATIAQLPNAELIIVTTPQAAADVAERAGSISLQTNOKVA  
 GVIENSGMVPDGSMEVFGSGGKIAQRISMLTGEDVPLVGLSVPLDPSLRGVGV  
 GNPIAISEPSPPTAAINAIAIDTLAKRKVSIIVKLTIGLVK"  
 complement(2513)..3271)  
 /note="CE1182"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAC17992.1"  
 /db\_xref="GI:23493021"  
 /translation="MTTATPWADAFDHVWDLFQILEVDHTERVAHLQBYFATDRT  
 GPMYTGSPNRLGGGDAQVNEITRDLISLSDVPLVGLVQLLEPPEDAQSI  
 DLEVVYDVTYVELPDLEQIRGLASIPDVLGSKAAEHLDAKNLWTLALRRNG  
 FETRVSKLIARAKPKLPIFIDSAIKQLNTNSGVFENFHTVLRAEDKALSRLKGI  
 RTERADQSGNALADISTIRVEDVVVWMDHKRT"  
 complement(3282)..3836)  
 /note="CE1183, similar to AE007003-3|AAK45527.1| percent  
 identity: 57 in 159 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="BAC17993.1"  
 /db\_xref="GI:23493022"  
 /translation="MADYSRSLDTPLVGTRRRRFKDDDTVGYAEKVARFQTQY  
 LFMQITFVIMLNTIGGFANWDPYPTLLNLAFTQAAVAPLILLAQNRDEDDK  
 STILADRRARETKADETELTRLAGVRLAVGEWTRDYLRHLELDLGRFLERLEAKL  
 DDAVARNDPHDLNAPTQKQVD"  
 complement(3829)..5124)  
 /note="CE1184, similar to AE007003-4|AAK45528.1| percent  
 identity: 52 in 417 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="BAC17994.1"  
 /db\_xref="GI:23493023"  
 /translation="MRGITRLYAGRLAGVVRGPTDVIIGVRVDRVVVTVIRSGGQDSRV  
 LGAVLVNNKRIIFLPMRLIAAIPKENVLISGSVLSARFKATGELQVTDIGITK  
 VHTDPELQHLGRAVELADVELETRTDWVLSVALIGERKFRGRTTYIAPWSH  
 WHGITAGGETNQTRELIAAGFEDMRADVAKELYGLPKPQDVVAGLELDERLADIL  
 QELSDRQAELETIERAADVLEMDPDADLIGELPEPKAGVILLEMDPEASAP  
 VRLINGFNPVTGALMTPEELTMDPSTTVAEALAMAKNPDLASLVFVVRPTAT  
 PTGKLGCVHLKREPPSLVGGILDPPLFYANDSQETAAARFFATYLVVCGPVL  
 DENGHLIGAVAVDLDLMDLPEDWRDAGIRPQVKEETLG"  
 5264..5782  
 /note="CE1185, similar to AE007003-6|AAK45530.1| percent  
 identity: 60 in 155 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="BAC17995.1"  
 /db\_xref="GI:23493024"  
 /translation="MSMQATGSGGNRAELGKKRPGVPGVSGFOTYBEAORAVDMUS  
 DNEPFAELTIVGDLMEVETVTRLTWGRVLGAGAAAGLGLFGLVGLGSLNLF  
 SLLTGLVGVVGVVLAAPVYVWASRGKRDFTSATQIVAGRYDILCRDPRAREARDMI  
 ARNMGLGHPRQ"  
 5805..6500  
 /note="CE1186"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAC17996.1"  
 /db\_xref="GI:23493025"  
 /translation="NRTGUKKALLGLVAVLGGSLVACGSDPNRPFPHQVAVQK  
 ENLVAVDDPIYGRNVISELYKTVLNGRGPVAVIRHEDNLIADPFLNKGTSILVIG  
 CTGALLEVNPVLEEELAAVAGETDINGEWRDRTFDALIGSIPRELLTADP  
 SNAQCFGANGLPQNVITYRNPFTFRADREILNWSGAITTSSELQELIEEAKELGRS  
 RELVKEFVASKG"  
 6660..7373

```

/notes="CE1187, similar to AB007262-3|AAK65408.1| percent
identity: 32 in 235 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17997.1"
/db_xref="GI:23493026"
/translation="METILIGVLLFVAACVQGAIGFGLMLAAPIALMRPDLPLGLI
LLAFGLSLATWARDREVENPVVWGLSLRIPGSLVGTWVAVMPVAGLKLIVATV
ILGLTSLVWSPGHGRNRSPLAGASGLGTTTGAIGPPLALIMRSMSPERVGTLS
VCFVIGSIALILTCAGALGMHLQALVYAPAVIAGVFLSGVWVKNLRLIFLGS
VFISLIGSNVITQAAGIF"
7592..8149
/notes="CE1188"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17998.1"
/db_xref="GI:23493027"
/translation="MGRGDFPAERPPADPRGKTIKMKWKGAIAAAAAALLV
GGGTFMSWDEVTQSDAIAAGNALSLADQKWTINGQDVNNVAVKVPVGDTLV
FTQOISVETSDMLSLTLTPGSITGATADAADALATELINDTITLIDGELPISIG
SGNTVTEGGVGVSGVSALSIKQAA"
complement(8190..9443)
/notes="CE1189, similar to AF024666-23|AAG03374.1| percent
identity: 84 in 408 aa"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAC17999.1"
/db_xref="GI:23493028"
/translation="MAAGPYSIDPTTVIDELLAQSPDLREMLQGFINOILSTQADQ
VGDYATVQSDRVNTRNGYRHRDTRVGTIVAVPKLRTGSPFDMLELRETRAER
ALUTTATCYLKGVSTRMNDLVATLGNLSKQSVSDMAKDLQWVEFRPRPDLTG
PLYVSCDALMKVRGGRVWTKSVLLATGNVAGRELLGQVATAESVAGWTFGR
DLKAGLEVLTSDAHGLOHAEVLPNASQRCRTHSKNLYMWSKOWTPLS
ANPHIFQPPDAQSVQAREVAFCEQKPPDVADYLEALDELAFNAKPSVTKV
WSNPTERINREIRRTDVGIFPNRDVAVRLVGVLAERQHDMLIQQRVMSLTSLEQ
TRAMTANIIDSDDITSEIRKQVA"
complement(11439..15143)
/gene="odha"
complement(11439..15143)

Query Match      34.5%; Score 818.6; DB 1; Length 308650;
Best Local Similarity 68.5%; Pred. No. 7.9e-190;
Matches 1182; Conservative 0; Mismatches 524; Indels 19; Gaps 3;

QY      650  TAACCATCAACATCAGTTTGATGGCCCAATGCGGTGCATCAACATCGCACGACGACGTTG 709
DB      190920  TAACCATCAGATCAGCTTACGGCCCAATCGGTGAGCACCAACGCGCACACCTATGTTG 190979

QY      710  ATCCAGCCCAACCTTGGGGCTGGAAGCGGGCTGAGCAATGCTGCGCGCGGAACCC 769
DB      190980  ATCCAGCCCAACCGCGGGCTGGAAGCGGGGACGCGACAGTGGCGCGCGCGTAAACG 191039

QY      770  ACCAGGGGAACAGATCAGCTTGCAGCGAAGCGCGGAGAAATCCACCGTCG 829
DB      191040  ACCAGTGGGAACAGATCAGCTTGCAGCGGAGCGGACCGCGAGGAGATCCACCGACCG 191099

QY      830  GTGTGCGCGTATTCGCGCGGACGCGCGCGGATAAAACAACAGCGTCCAAATACGATC 889
DB      191100  GTCTCCCGGTACTGGGCTCGACACCAACCGATGAGACGAAGGATCCAGGTAGGCATTG 191159

QY      890  GGGTTCACACAGGTGACGATGCGATCAACATGGCTTTTACCACAAACCCGCTGTTA 949
DB      191160  GGATTGAGCGAGTCAGACATGGCCATGAGCATGGGCTGACCCAGACCTCCCGGCTG 191219

QY      950  TCAGCGCTCACCTCCACCGCACCGGTTGCGGTGTCAGTGGCCACCGCGCAACCGGCC 1009
DB      191220  CCTGATGATGAGCGAGCGCGGGTCTGTTGTTGTCGTCACGCCCCCGCGGAGCGGAC 191279

QY      1010  AAGAGCGTGTGATCGGCGACGTTGGTTCTGTTTCTTCAATGATGTGGGCGCTTCACCC 1069

```

```

191280  GCTGCGC---CAACGGGTTTGGAAATGCTC-----GACAAAGTTTACCTCGGTG 191324
QY      1070  TTCTTTGTCTAGCGGTCTTTTCGCTCCATGACGCGCAAAACATPAACAGGTAGGATGCCA 1129
DB      191325  CGGGCGGACGGGCTGCGCGCGCCCATCACCGCAACCAACAGCAGGTAGGCGGATGCGG 191384
QY      1130  CCCAGCGCATATATATCGAGCAGATCGGCGCGCATTTGGACAAAGATCAACGCCCAAG 1189
DB      191385  CACACGCGCAGATGTCGAAATGATCGGGCGGTGTCCGAGATTCGCGCCGACCCCGAGG 191444
QY      1190  GTCCCGGATGAAACAAAAGAGCTCAGAAATTAACACACACGAGAAAGACCGCAATGAT 1249
DB      191445  GTCCCGGAGGTGAACAGACCACTCGGACGACGACGACGACGATGATCAGCGCGGTGATG 191504
QY      1250  CCTTTCGCGCTTAATTCCTTTTAAATCAACAGTACATCTGCGGTCCGATGGACAGTAA 1309
DB      191505  CCTTCGCGTTTGTATGCTCTGTTTGTATCACCAGGACATTTCTGTGGGCGCATGGCCAC 191564
QY      1310  AGACTGCGCCCAAAAGCAGACCTGTAATGAAGATTTCCATGATCACCATCGTACCTAT 1369
DB      191565  AGACTGCGTCCCAACCAACACCGGTGACAAAGATTTCCATGTCGCGCATGTCGCGCAC 191624
QY      1370  GGAAGTACTTAAGTAAATGATTTCTTAACATGTTTAAATAGCTTCAATGATGATGAC 1429
DB      191625  AAACCATTAAGTAAATGATGAGACTGAACATGCTTTAGAAATGCTTTCATGACCCG 191684
QY      1430  ATTCACCTGACACTTTGCTCTCAATCATGATGAAGCAGCTTCGAAGGCGCTCTCTTA 1489
DB      191685  ATTCACCTGACACTCTCTGACCATCATCGACGAGGCGAGCTTCGAGAACGCTCTCTC 191744
QY      1490  GCGCTTTCATTTCCCGCTCGGCGGTGAGTCAGCGCTTAAAGCTCTCGAGATCAACGTG 1549
DB      191745  GCATCTGCTGATTCCTCCCTTCGCGGTGAGTCAGCGCTTAAAGCTCTCGAGATCAACGTG 191804
QY      1550  GGTGAGTGTGTTGATTCGCGCACCAACCGGCGCAAGCAACCGAGCGGCTCAAGTCTTT 1609
DB      191805  GGTGCGGTACTGTGTGCGCACCAACCGGCGGTGTCACCGAGCGGCGGAGGTGCTG 191864
QY      1610  GTGCAAGCGCGCGGAAATGTTGTTCTGACGAGCAAACTAAAGCGCACTATCTGGA 1669
DB      191865  GTCCAGCGCGCGCGCAAGATGCGCTGCTGAGGCGGAGACCCGTGAGCACTGCGCGAA 191924
QY      1670  CGCTTCTGAAATCCCGTTAACCATCGCATCAAGCAGATTCGCTTATCCATGCTTT 1729
DB      191925  CGGCTTGACAGATTCCTCCCTGACCTGCTATCAACGCGGATTCCTGTCACCTGCTTC 191984
QY      1730  CTCTCCGTTTCAACAGAGTGTGTTGGGTGAGAGCAAGCTCAGCTGCGCTTGGAA 1789
DB      191985  CCGCCCGTGTTCGCGGAGGTTCACATTTGGGGTGGCGTCACTCCCTGACGCTGCTGGAG 192044
QY      1790  GATGAAGCGCACATTTCTCTGCGCGGTGAGAGATGTTTAGGAGCGGTAAACCGCT 1849
DB      192045  GACGAGGCGCACACCTCTGCTGTCGCGCGGTTCGCTGTCGCGCGGTTCGCTGCGCGCG 192104
QY      1850  GAAGCTAATCCCGTGGCGGATGTGAAGTAGTAGAACTTTGGAACCATGCGCGCATTTGGCC 1909
DB      192105  GAGCGGACCCCGTGTGCGGGTGTGAGGTCTCTACGCTCGGGGTTCATGAGACCTGCGG 192164
QY      1910  ATTGCAACCCCTCATTTGCGGAGATGCTACTATGTTTGTGTGGAACTAGATTTGGGCTGG 1969
DB      192165  GTGGCCACCCCGGAACCTGAGGCGACGCTACACCGTGTGATGGCCAAACCCGACTGGGTG 192224
QY      1970  ATGCGCGTCTTACGCTTCCGCTCCCAAGATGCTGCTTCAAGACCGGTGACCTGGAACGG 2029
DB      192225  ATGCGGTTACTGCGTTTTCGCGCCGATGATGCTGTGTCAGACCCGCGATCTCGAGGCG 192284
QY      2030  GTGAGGTCTCTGTGGGCGCGACGCGGTATTCATGTTCCGTTCGCGCGAGGTGTTGCT 2089
DB      192285  GTGACGAGCGCTGCGCACGCGCGGTGCTGCGGTGCGCGCTCTGCGGAGGTTTCGGG 192344
QY      2090  GAGGCAATTCGCGGAGCGCTTGGTTCGGGACTTCTTCCCGMAACCCAAAGTGTCTCCATG 2149
DB      192345  GAGGCGGTGCGACTCGCGCTGGGGTGGGGCTGCTGCTCCCGAGGCCCGCGCGCGATG 192404

```



Db 790 GGTGGGTACTGTGTGCGGCACCCAGCCGCGGTTGTCACCGAAGCCGCGGAGTGCTG 731  
 QY 1610 GTGCAAGCAGCGGGAATATGTGTCTGCAAGCAGAACTTAAAGCGCAACTTCTCGA 1669  
 Db 730 GTCCAGCGCGCCGAGATGCGGTCTGCAAGCGGAGACCGGTGAGCAACTGCGGAA 671  
 QY 1670 CGCTTGTCTGAATCCCGTTAAACATGCGCAAGATGCTATCCACATGCTTT 1729  
 Db 670 CGCTTGTGAGAGATCCCGTCTGATCAAGCGCGATTCCTGTCCACCTGGTTC 611  
 QY 1730 CTTCCGCTGTTCAAGCAGGTAGCTTCTGGGCTGGAGCAACGCTCAGCTGCGCTGGAA 1789  
 Db 610 CCGCCCGTGTTCGCGGAGGTGCAATTTGGGCTGGTCACTGACGCTGCGTGGAG 551  
 QY 1790 GATGAAGCGCACATTTATCTTCTGCTGCGCGGTGGAGATGTTTTAGAGCGGTAAACCGGT 1849  
 Db 550 GACGAGCGCCACACCTGTCTCTGTTGCGCGGCTGCTGCGGCGGTGACCCGC 491  
 QY 1850 GAAGCTAATCCGCTGGCGGATGTGAAGTAGTAGAATTTGGAACCATGCGCCACTTGGCC 1909  
 Db 490 GAGCGGACCCCGTGTGCGGGTGTGAGGTCTTACGCTGCGGGTCTATGACACCTTGGC 431  
 QY 1910 ATTGCAACCCCTCATTTGCGGATGCTACATGCTTGTGATGGGAAACTAGATTGGCTGCG 1969  
 Db 430 GTGGCCACCCCGGAACTGAGGCGACGTTACACCGTGTGATGGCCACCGACTGGGTGCG 371  
 QY 1970 ATGCCGCTTTACGCTTGGTTCCTCAAGATGCTTCAAGCGGTGATGCGGCTGAGCGGCGC 2029  
 Db 370 ATGCCGCTTCTGCTTTCGCGCGGATGATGCTGTCAGGACCGCGATCTCGAGGGGAGG 311  
 QY 2030 GTGATGCTCTGCGGCGCGGCGGTATCTTCTGCTGCGCGGAAAGTGTGGT 2089  
 Db 310 GTCCAGGAGCGGCTGCCAGCGCGGCTGCTGCTGCTGCGGAGGTTTCGAG 251  
 QY 2090 GAGCAATTCGCGGAGCGCTTGTGTGGGACTTCTCCGAAACCCAAAGCTGCTCCCATG 2149  
 Db 250 GAGCGGTGCGACTCGGCTGGGTGGGCTGCTGCGGCGCGCGCGCGCGATG 191  
 QY 2150 CTAAGCAGGAGAGTATCTCTGCTGATGATACCCATGACACCGCATCTATTGG 2209  
 Db 190 CTGGCTGCGCGGATGTGGTGCAGTGGATGAGAGGTGGTGGACACCCCGCTGTATTGG 131  
 QY 2210 CAAGATGCGGCTGGAATCTAGATCTCTAGCTAGCTACAGACCGCTGTTGATGCA 2269  
 Db 130 CAAGATGCGGCTGGAATCTAGATCTCTAGCTAGCTACAGACCGCTGTTGATGCA 71  
 QY 2270 GCAATTCGAGGATGCGGCTTGTATCTTCTGAAAGGTTCAGGCTTTTCACTTCTTC 2329  
 Db 70 GCGCGGCGGCTGAG----AACCTAACTGGAACGTTCCGGGTTCTTGACCTCTTC 15  
 QY 2330 GCCCGCAGGAATG 2343  
 Db 14 ACCGAGGATGCG 1

RESULT 13

EX248357 349535 bp DNA linear BCT 06-NOV-2003  
 LOCUS Corynebacterium diphtheriae gravis NC131129, complete genome;  
 DEFINITION segment 4/8.  
 ACCESSION BX248357  
 VERSION BX248357.1 GI:38199912  
 KEYWORDS complete genome.  
 SOURCE Corynebacterium diphtheriae  
 ORGANISM Corynebacterium diphtheriae  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
 REFERENCE 1 (bases 1 to 349535)  
 AUTHORS Cerdano-Tarraga A.M., Efstathiou A., Dover L.G., Holden M.T.G.,  
 Pallen M., Bentley S.D., Besra G.S., Churcher C., James K.D., De  
 Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,  
 Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,  
 Rabinowitz E., Rutherford K., Thomson N.R., Unwin L.,

TITLE Whitehead S. and Barrell B.G. Parkhill, J.  
 JOURNAL The complete genome sequence and analysis of Corynebacterium  
 PUBLISHED diphtheriae NC131129  
 REFERENCE Nucleic Acids Res. 31 (22), 6516-6523 (2003)  
 AUTHORS 14602910  
 DIRECT SUBMISSION 2 (bases 1 to 349535)  
 TITLE Cerdano-Tarraga A.M.  
 JOURNAL Direct Submission  
 Submitted (03-OCT-2003) Cerdano-Tarraga A.M., submitted on behalf  
 of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust  
 Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:  
 amct@sanger.ac.uk  
 FEATURES  
 Location/Qualifiers  
 1. 349535  
 /organism="Corynebacterium diphtheriae"  
 /mol\_type="genomic DNA"  
 /strain="NC131129"  
 /db\_xref="taxon:1717"  
 /note="biotype gravis"  
 76. 402  
 /locus\_tag="DIP1058"  
 76. 402  
 /locus\_tag="DIP1058"  
 /note="Similar to Mycobacterium tuberculosis hypothetical  
 11.0 kDa protein Rv1898 or MT1949 or MTCV180.20c  
 SW:Y198 MYCTU (007734) (102 aa) fasta scores: E():  
 5.8e-16, 56.43% id in 101 aa, and to Aquifex aeolicus  
 hypothetical 17.9 kDa protein AQ.2067 TR:067847  
 (EMBL:AB000771) (157 aa) fasta scores: E(): 4.7e-08,  
 40.19% id in 102 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Conserved hypothetical protein"  
 /protein\_id="CAE49578.1"  
 /db\_xref="GI:38199913"  
 /translation="MLDMLAFVAPPTENKAKMADVAEIRVRESGLPNETN  
 AMFLIEGSEWMAVVKATDAVLAVSPRAVLKADIRPGYTGLOQKVASVERVL  
 AGDSGS"  
 88. 372  
 /locus\_tag="DIP1058"  
 /note="HMPfam hit to PF01910, Protein of unknown function  
 DUF77"  
 complement (405. 4762)  
 /note="Putative iron uptake system. Low GC content  
 (58.02%)"  
 complement (405. 1226)  
 /locus\_tag="DIP1059"  
 complement (405. 1226)  
 /locus\_tag="DIP1059"  
 /note="Similar to Streptomyces coelicolor putative  
 iron-siderophore uptake system ATP-binding component  
 SC151.25C TR:09S215 (EMBL:ALU09848) (301 aa) fasta scores:  
 E(): 1.9e-42, 50.58% id in 255 aa, and to Escherichia coli  
 ferric enterobactin transport ATP-binding protein fepC  
 b0588 SW:PEPC.BCOLI (P23878) (271 aa) fasta scores: E():  
 1.1e-41, 49.61% id in 262 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Putative iron-siderophore uptake system  
 ATP-binding component"  
 /protein\_id="CAE49580.1"  
 /db\_xref="GI:38199914"  
 /translation="WTSTPILPDSVTLWDKHTVSTDLTVRPGQFTAIIGPNC  
 GKSTLLKSFARILQPHQTQVTLGPNLDSLTHTKVIQRIALLPQALAPADITVEELV  
 GRGPFYQTLKQWSTDAQVDSALLATGTAQRHRTVELSGGQRQVWLMVLAQ  
 NTPVVLDEPTTLDVAHQVLELALRQQLGRTIVTVLHDLQQAQVADHLIWMK  
 GGAIVAGAPSSIMTPQLISSEFINDVHTTIDHVVIVKTPQPDVSVLENLQ"  
 complement (564. 1133)  
 /locus\_tag="DIP1059"  
 /note="HMMsmart hit to SM00382, ATPases associated with a  
 variety of cellular activities"  
 complement (567. 1130)  
 /locus\_tag="DIP1059"  
 misc\_feature  
 misc\_feature  
 gene  
 CDS  
 misc\_feature  
 misc\_feature  
 gene  
 CDS  
 misc\_feature  
 misc\_feature

```

/misc_feature
/note="HMPfam hit to PF00005, ABC transporter"
complement(579..797)
/locus_tag="DIP1059"
/note="ProfileScan hit to PS0100, 2nd half motif for
nucleotide binding, associated with P-loop."
complement(753..797)
/misc_feature
/locus_tag="DIP1059"
/note="ScanRegExp hit to PS00211, ABC transporters family
signature."
complement(1047..1124)
/misc_feature
/locus_tag="DIP1059"
/note="ProfileScan hit to PS0101, P-loop nucleotide
binding motif (does not find all)."
complement(1086..1109)
/misc_feature
/locus_tag="DIP1059"
/note="ScanRegExp hit to PS00017, ATP/GTP-binding site
motif A (P-loop)."
complement(1223..2182)
/misc_feature
/locus_tag="DIP1060"
complement(1223..2182)
/locus_tag="DIP1060"
/note="Similar to Streptomyces coelicolor putative
iron-siderophore uptake system transmembrane component
sc151.26C TR:Q9S214 (EMBL:AL109848) (375 aa) fasta scores:
E(): 5e-40, 40.74% id in 324 aa, and to Escherichia coli
ferric enterobactin transport system permease protein fepg
fep9 or B0589 SW:FEFG_ECOLI (P23877) (330 aa) fasta
scores: E(): 1.8e-45, 44.09% id in 322 aa"
/codon_start=1
/transl_table=11
/product="Putative iron-siderophore uptake system
transmembrane component"
/protein_id="CAE49582.1"
/db_xref="GI:38199915"
/translation="NRALVLSLLVLTAAVSLVPGAGLSTWQLTSETPLATVVM
QRLPRVTVGLVGAALAGLSFQSLTRPLGSPDIFGSTGAVTGVIAPFLGWSG
FGATVIGLGLVAVVVALSVTRIDGLRIILVGLIGLSMLGALNRMLITRGELD
TALSASWAGSLGLRWTAIPACCVLAITITPLRFDLVSLGDLVAGLGLR
LQLKLLLVGVFLVAVTAVAGPIAFVALASPHARLTSRTPLVENSIVGALL
LIVADLLGRLRFQLFVGLVTVTIGMYLLNLVLRGSKENS"
complement(2108..2182)
/misc_feature
/locus_tag="DIP1060"
/note="Signal peptide predicted for DIP1060 by SignalP 2.0
HMM (Signal peptide probability 0.966) with cleavage site
probability 0.180 between residues 25 and 26;
signal-peptide site"
complement(1244..2116)
/misc_feature
/locus_tag="DIP1060"
/note="HMPfam hit to PF01032, FecCD transport family"
order(complement(2105..2161), complement(1979..2044),
complement(1883..1939), complement(1805..1870),
complement(1718..1783), complement(1607..1657),
complement(1445..1510), complement(1319..1384),
complement(1247..1303))
/misc_feature
/locus_tag="DIP1060"
/note="Tg probable transmembrane helices predicted for
DIP1060 by TMHMM2.0"
complement(1250..1564)
/misc_feature
/locus_tag="DIP1060"
/note="BlastProDom hit to PD001557, PD001557"
complement(2184..3191)
/misc_feature
/locus_tag="DIP1061"
complement(2184..3191)
/locus_tag="DIP1061"
/note="Similar to Streptomyces coelicolor putative
iron-siderophore uptake system transmembrane component
sc151.27C TR:Q9S213 (EMBL:AL109848) (348 aa) fasta scores:
E(): 8e-36, 36.53% id in 323 aa, and to Escherichia coli
ferric enterobactin transport system permease protein fepd
or B0590 SW:FEPD_ECOLI (P23876) (334 aa) fasta scores:
E(): 5.2e-40, 40.37% id in 322 aa"
/codon_start=1
/transl_table=11
/product="Putative iron-siderophore uptake system
transmembrane component"
/protein_id="CAE49584.1"
/db_xref="GI:38199916"
/translation="MLRQPRVRLPPLVITLVIILCTIIMSLGVGARNITSPDILA
ALRHGYDPTNSHITWGRIPRTILIALAGASISLAGVLIOALTREPLADTGVGINAG
AAFTVIGLGLSHISIFALAILAGALGACVYALSNWCKSGDFRLVLAGVAL
SAULTIGLGLVNPQAFDRKSNMGNIDAGSYQPAAVAGVGLILGLVATCMHQ
LNLISLDELATWGAISTRLFTIATVIAASATAAAGVITFLGLMVPVIAWTV
GPNLLRLVATSLGIPIVLSDILGRIIVPGEPFAGVWVAFIAPFLIAYQTKKE
I"
complement(3090..3191)
/misc_feature
/locus_tag="DIP1061"
/note="Signal peptide predicted for DIP1061 by SignalP 2.0
HMM (Signal peptide probability 1.000) with cleavage site
probability 0.663 between residues 34 and 35;
signal-peptide site"
complement(2199..2525)
/misc_feature
/locus_tag="DIP1061"
/note="BlastProDom hit to PD001557, PD001557"
complement(2205..3095)
/misc_feature
/locus_tag="DIP1061"
/note="HMPfam hit to PF01032, FecCD transport family"
order(complement(3096..3161), complement(2934..2990),
complement(2838..2903), complement(2766..2831),
complement(2670..2735), complement(2544..2609),
complement(2394..2459), complement(2289..2354),
complement(2211..2276))
/misc_feature
/locus_tag="DIP1061"
/note="Tg probable transmembrane helices predicted for
DIP1061 by TMHMM2.0"
complement(3239..4282)
/misc_feature
/locus_tag="DIP1062"
complement(3239..4282)
/locus_tag="DIP1062"
/note="Similar to Escherichia coli iron(III)
dicitrate-binding periplasmic protein precursor FecB or
B4290 SW:FECB_ECOLI (P15028) (300 aa) fasta scores: E():
Query Match 30.5%; Score 724.4; DB 1; Length 349535;
Best Local Similarity 58.2%; Pred. No. 1.2e-166;
Matches 1351; Conservative 0; Mismatches 931; Indels 38; Gaps 3;
QY 1 AGATACCTCTTTGGAAGAACCAATGACGATTCGCGTAGCATTTGTCGCTGGAAGGC 60
DB 35950 AGATACGCTTTGGAAGAAATGCATATGCTTTGCGAGACATCGTGGCATCCGCAAGGC 36009
QY 61 TCTTTAGCTGGGTATTTCTTCAGTCCAGAGCTCAGACGGAGCGGCTGAGTTTCA 120
DB 36010 TCCTATGTGGGAATTTCTCTTATGGACCTGAATCAACCGAGCTATCTATTTGAACCGCTG 36069
QY 121 GCGGAGGAGGCTGCCCGCTTCTGATTTCATCAGCCAAAGCTATTCATCATTAATCGTTG 180
DB 36070 AGCTGTGTAGGCTGCCCGCTGCTGATTTCATCAACCGAGCTATCTATTTGAACCGCTG 36129
QY 181 GGTGGAGAACCCGGCGGATGACGGTGAGAACTTTGTCAGTCAGTCGCAACATGCT 240
DB 36130 GGTAGAAGAACCGGTGAAGATGCGGAGAACCTGTAGATCTGCGGGCAATAACGACT 36189
QY 241 TGGCGTCATGCTTTCTCACCCTTCGCGCAGGCGCTCTCAGCGCAAAATATCTCCATGG 300
DB 36190 GGGTGTATTGATTTTGGCACTTTCACAGGGGCTTTTAACTGACAGGTATCTCGAGG 36249
QY 301 AATTCCAGAGGGTTCGCCGCCAGCCAGCGGTAACTCCCTGTCTGAGGCGCATGTTGAACGT 360
DB 36250 CGTGCCAGTGGATTCACGGCAGCAGCAGGGAAGTCTCTGGGCAAGAGATGCTCAACGC 36309
QY 361 GAACAATATTGATATGTCGCGAAGCTCAATGACATGCCCGAGGACGGGCGAGTCACCT 420
DB 36310 TAAAAATCTTGACATGGTGGCTGCACTGAATGATTTTAGCGTTACGACGTCGCCAAGCTI 36369
QY 421 TGGCGAGTGGCGCTTTCATGGGTCTGCGCGCAGCAAGGAGATACGCGCGGATACCGT 480

```



1535	CTCGAGCATCATGNGGGTCTGAGTGTGGTATTCGCGCACCCCAACCGGCCAAGCAACCGAA	1594	
Db	37498	TTGGAGCGGGAGACATGGTCTGAGTGTAGTGTGACCGGTACGTCCACAGTCACTGCGAACTGAT	37557
Qy	1595	GCGGGTGAAGTCTTTGTGCAAGACGCGCGGAATAATGGTGTCTGCAAGCAGAAACTAAA	1654
Db	37558	GCGCGGAGGTACTGGTGTGCAATCGGCACGACGTATGGCACTTTCTGCAAGGCCGGAACAAAC	37617
Qy	1655	GCGCACTATCTGACGCGCTTCTGAAATCCCGTTAAACCATCGCCATCAACGCAAGATTGG	1714
Db	37618	GCTCAGTCTCGCGGGAGAAATTGAGCGGAATTCATTTCTGTGTAGCTGTCAACGCAAGATCT	37677
Qy	1715	CTATCCACATAGTTTCTCCCGTGTTCACAGAGTAGTCTTTGGGGTGGAGCAACGCCTC	1774
Db	37678	TTGGCCACATGTTTTCGCCCGGTGATGGCAAGTGTGGCGGAGATGGATAAGCCGCACTTG	37737
Qy	1775	ACGTGCGCTTGGAAAGATGAAGCGGCACATTAATCTTTGCTGCGCGTGGAGATGTTTATA	1834
Db	37738	CATATGCGTATCGAGGACGAGTGCATTCGTTTCGTTGTGTGGTTCGGGAGATGTCCTT	37797
Qy	1835	GGAGCGGTAAACCGCTGAAGCTAAATCCCGTGGCGGGATGTGAAGTAGTAGAACTTGGAAAC	1894
Db	37798	GGCGCTGTAAACGGGGGAAAAGACACCTGTATCTGGTTGCCACTCGATAGCTTAGGTGCA	37857
Qy	1895	ATGCGCCACTTGGCCCATTTGCAACCCCTCATTTGCGGATGCCCTACATGCTTGATGGGAAA	1954
Db	37858	TTCCGCATATTTTGCAGTGGCAAGTCCGAAGCTGTTCAGCTTTCGATCGATAGCTTAGGTGCA	37917
Qy	1955	CTAGATTGGGCTGCAGATCCCGTCTTACGCTTTCGATCCCAAGATGTGCTTCAAGACCGT	2014
Db	37918	GTGTATTTGGAGCGGATGCGCAGTATTCGATTTGGCCCTCGGATGGGCTTCAAGACCGT	37977
Qy	2015	GACCTGGACGGCGCGTCGATGTGCTGTGTGGGCGCAGCGCCTATCCATTTGTCCTCGCTCG	2074
Db	37978	GATATTAGCGTCTGTTTGGTGAAGTTCGCGCAGGGCGAGTGTGTAAAGAGATCCCTTCA	38037
Qy	2075	GCGGAAGTTTTGGTCTAGGCAATTCGCCGAGGCGCTTGGTTGGGGACCTTTTCCCGAAACC	2134
Db	38038	TCGGAGGCAATTCATGAGACGGCCCGTGTGGGATTTGGGCTGGGCAATGCTCCCGCGATTT	38097
Qy	2135	CAAGCTGCTCCCATGCTAAAGCAGGAGAAAGTATCCTCTTCGATCGATACCATCCATTGAC	2194
Db	38098	CAAGCGTTACCGCTTTTGGATTTCTGCTGAGTAGTGTATGCTAGATAGCGAGGTGCACGAA	38157
Qy	2195	ACACCATGATTTGGCAAGATGGCGCTCGGAATCTAGATCTTAGCTAGACTCACAGAC	2254
Db	38158	GTTTCGCTGTATTGGCAGCGATGGCGGTTCGATCTCCAGGCCCTAGAGCGGTGACTCAA	38217
Qy	2255	GCGCTCGTTGATGCAGCAATCGAGGGATTTGGCGCTTAGT	2294
Db	38218	GCTGTAGTGTATGCAGCGCTCCGCTTCGAGGGAATCACTACCACTGT	38257

RESULT 14						PAT 24-FEB-2003
AX643028/c						
LOCUS			711 bp	DNA		linear
DEFINITION			Sequence 7 from Patent EP126966.			
ACCESSION			AX643028			
VERSION			AX643028.1 GI:28550158			
KEYWORDS			Corynebacterium glutamicum			
SOURCE			Corynebacterium glutamicum			
ORGANISM			Bacteria; Actinobacteria; Actinomycetales; Coryneбактерiaceae; Corynebacterium.  			
REFERENCE AUTHORS		Gunji,Y. and Yasueda,H.				
TITLE		Method for producing L-lysine or l-arginine by using methanol assimilating bacterium				
JOURNAL		Patent :EP 126966-A 7 18-DEC-2002;				
Ajinomoto Co., Inc. (JP)						
Location/Qualifiers						
FEATURES						
source						1..711



CDs

/organism="Corynebacterium glutamicum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1718"  
1. 711  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAD67811.1"  
/db\_xref="GI:28550159"  
/translation="MVIMEIFITGILGASLLSIFQNVLVIKQIKREGLIAVLV  
CLISDFPLFTGDLVLLSNAAPIVLDIMWGGIAYLLFWFAWAAKDMNVEAPQ  
IIBETPTVDDTPLGSAVATDNRVRSVSKORVWVPMMAVLVTLNPNAY  
LDAPFIVGGVQAGDGRWIPAGAPASLLWPLVPGFAGAAALSRPLSSPKVWRIN  
VVAVVMTALAKMLMG"

ORIGIN		Query Match	Score	711;	DB 6;	Length	711;
		Best Local Similarity		100.0%;		Pred. No. 1.2e-163;	
		Matches		711;		Conservative	
		0;		Mismatches		0;	
		Indels		0;		Gaps	
QY	649	CTAACCCATCAACATCAGTTT	GATGGCCAAATGCGT	CTATCAAACTGCCACGACGCTT	708		
DB	711	CTAACCCATCAACATCAGTTT	GATGGCCAAATGCGT	CTATCAAACTGCCACGACGCTT	652		
QY	709	GATCAGCGCCACACCTTGGGCT	GGACAGCGGGGTGACATGCT	GCTGCGCGAAACC	768		
DB	651	GATCAGCGCCACACCTTGGGCT	GGACAGCGGGGTGACATGCT	GCTGCGCGAAACC	592		
QY	769	CACACGCGGAAACAGATCAGGCT	TTGCCGGAACGCGCGCGGCGGAAAT	TCCACCGTCC	828		
DB	591	CACACGCGGAAACAGATCAGGCT	TTGCCGGAACGCGCGCGGCGGAAAT	TCCACCGTCC	532		
QY	829	GGTGTGCGCGTATTTGGCGCG	CGACCGCGCGATTAACAACAAGCGGT	CGAAATACGAT	888		
DB	531	GGTGTGCGCGTATTTGGCGCG	CGACCGCGCGATTAACAACAAGCGGT	CGAAATACGAT	472		
QY	889	CGGGTTCAACAGGTGACGAGT	TGCGCATCAATGGGCTTTACCCAAACCGCTGCT	948			
DB	471	CGGGTTCAACAGGTGACGAGT	TGCGCATCAATGGGCTTTACCCAAACCGCTGCT	412			
QY	949	ATCGAGCTCACTTCAACCGCA	CCGCGTTCAGTGGCCACCGCGCAACCGCC	1008			
DB	411	ATCGAGCTCACTTCAACCGCA	CCGCGTTCAGTGGCCACCGCGCAACCGCC	352			
QY	1009	CAAGCGGTGTCATCGGCGCA	CGGTTCGTTCTTCAATGATCTGGCGCTTCCAC	1068			
DB	351	CAAGCGGTGTCATCGGCGCA	CGGTTCGTTCTTCAATGATCTGGCGCTTCCAC	292			
QY	1069	CTTTGTTGTCATGGGCTCTT	TGCGTCCATGAGCGAAACCAATACAGGTATGCG	1128			
DB	291	CTTTGTTGTCATGGGCTCTT	TGCGTCCATGAGCGAAACCAATACAGGTATGCG	232			
QY	1129	ACCCAGCGCATATATCGAGCA	CGATCGGCGGCGATTGGAACAAAGATCAACCGCCAA	1188			
DB	231	ACCCAGCGCATATATCGAGCA	CGATCGGCGGCGATTGGAACAAAGATCAACCGCCAA	172			
QY	1189	GGTGGCGCGATGACAAAAG	AGGTGAGAAATTAACACACGAGAAACCGCAATGAG	1248			
DB	171	GGTGGCGCGATGACAAAAG	AGGTGAGAAATTAACACACGAGAAACCGCAATGAG	112			
QY	1249	TCCTTCGCGCTTAATCTCT	TTGTTAATCACCAGTACATCTCGCGTCCGATGGACAGTAA	1308			
DB	111	TCCTTCGCGCTTAATCTCT	TTGTTAATCACCAGTACATCTCGCGTCCGATGGACAGTAA	52			
QY	1309	AAGACTGGCCCCAAAAGCAG	ACCTGTAAATGAAGATTTCATGATCAACAT	1359			
DB	51	AAGACTGGCCCCAAAAGCAG	ACCTGTAAATGAAGATTTCATGATCAACAT	1			

RESULT 15  
AX123539/c  
LOCUS

Sequence 3455 from Patent EP1108790.  
AX123539  
AX123539.1 GI:14041027  
Corynebacterium glutamicum  
Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriaceae; Corynebacteriidae; Corynebacterium.  
1  
Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.  
Novel polynucleotides  
Patent: EP 1108790-A 3455 20-JUN-2001;  
KYOMA HAKKO KOGYO CO., LTD. (JP)  
Location/Qualifiers  
1. 708  
/organism="Corynebacterium glutamicum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1718"

ORIGIN		Query Match	Score	708;	DB 6;	Length	708;
		Best Local Similarity		100.0%;		Pred. No. 6.5e-163;	
		Matches		708;		Conservative	
		0;		Mismatches		0;	
		Indels		0;		Gaps	
QY	652	ACCCATCAACATCAGTTT	GATGGCCAAATGCGT	CTATCAAACTGCCACGACGCTT	711		
DB	708	ACCCATCAACATCAGTTT	GATGGCCAAATGCGT	CTATCAAACTGCCACGACGCTT	649		
QY	712	CGACCGCCACACCTTGGGCT	GGACAGCGGGGTGACATGCT	GCTGCGCGAAACCCAC	771		
DB	648	CGACCGCCACACCTTGGGCT	GGACAGCGGGGTGACATGCT	GCTGCGCGAAACCCAC	589		
QY	772	CAGCGGAAACAGATCAGGCT	TGCCGCGAAGCGCGCGCGGCGGAAAT	TCCACCGTCCGT	831		
DB	598	CAGCGGAAACAGATCAGGCT	TGCCGCGAAGCGCGCGCGGCGGAAAT	TCCACCGTCCGT	529		
QY	832	GTCGCGGTATTTGGCGCG	CGACCGCGGATAAACAACAAGCGGT	CGAAATACGATTCGG	891		
DB	528	GTCGCGGTATTTGGCGCG	CGACCGCGGATAAACAACAAGCGGT	CGAAATACGATTCGG	469		
QY	892	GTTCAACAGGTGACGACG	ATGCGCATCAATGGGCTTTACCCAAACCGCTCTTATC	951			
DB	468	GTTCAACAGGTGACGACG	ATGCGCATCAATGGGCTTTACCCAAACCGCTCTTATC	409			
QY	952	GAGCTCACTCCACCGCG	CGACCGGTTCAGTGGCCACCGCGCAACCGCCAA	1011			
DB	408	GAGCTCACTCCACCGCG	CGACCGGTTCAGTGGCCACCGCGCAACCGCCAA	349			
QY	1012	AGGCGTGTCAATCGGCG	ACGCTGCTTCTTCAATGATCTGTGGCGCTTCCACCTT	1071			
DB	348	AGGCGTGTCAATCGGCG	ACGCTGCTTCTTCAATGATCTGTGGCGCTTCCACCTT	289			
QY	1072	GTTTGTGTCATGGGCTCTT	TGCGTCCATGAGCGAAACCAATACAGGTAGCGATGCCACC	1131			
DB	288	GTTTGTGTCATGGGCTCTT	TGCGTCCATGAGCGAAACCAATACAGGTAGCGATGCCACC	229			
QY	1132	CCAGCGCATATATCGAC	ACGATCGCGCGGCGATTGGACAAAAGATCAACCGCCAAAGGT	1191			
DB	228	CCAGCGCATATATCGAC	ACGATCGCGCGGCGATTGGACAAAAGATCAACCGCCAAAGGT	169			
QY	1192	CCCGCGCATGACAAAAG	AGCGTCAGAAATTAACACACGAGGAGGACCGCAATGATGCC	1251			
DB	168	CCCGCGCATGACAAAAG	AGCGTCAGAAATTAACACACGAGGAGGACCGCAATGATGCC	109			
QY	1252	TTCCGCGTAAATCTCT	TTTATACACAGTACATTTCTGGGCTCCGATGGACAGTAAAG	1311			
DB	108	TTCCGCGTAAATCTCT	TTTATACACAGTACATTTCTGGGCTCCGATGGACAGTAAAG	49			
QY	1312	ACTGGCCCCCAAAGCAG	ACCTGTAAATGAAGATTTCATGATCAACAT	1359			
DB	48	ACTGGCCCCCAAAGCAG	ACCTGTAAATGAAGATTTCATGATCAACAT	1			

Thu Mar 18 12:31:04 2004

us-09-105-117k-3.rge

Page 20

Search completed: March 15, 2004, 22:02:11  
Job time : 9512 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 15, 2004, 10:09:44 ; Search time 968.046 Seconds  
(without alignments)  
10418.122 Million cell updates/sec

Title: US-09-105-117K-3  
Perfect score: 2374  
Sequence: 1 agatactccttgaagaaa.....gtacaccttcagcaaatgg 2374

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124093041 residues  
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 25Jan04:\*  
1: Geneseqm1980s:\*  
2: Geneseqm1990s:\*  
3: Geneseqm2000s:\*  
4: Geneseqm2001as:\*  
5: Geneseqm2001bs:\*  
6: Geneseqm2002s:\*  
7: Geneseqm2003as:\*  
8: Geneseqm2003bs:\*  
9: Geneseqm2003cs:\*  
10: Geneseqm2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	2374	100.0	2374	2 AAT96816	Aat96816 DNA encod
C 2	2374	100.0	2374	9 ADB66196	Adh66196 DNA fragm
C 3	2374	100.0	349980	5 AAH68528	Aah68528 C glutami
C 4	993	41.8	993	4 AAF71777	Aaf71777 Coryneb
C 5	993	41.8	993	4 AAS96096	Aas96096 C. glutam
C 6	870	36.6	870	5 AAH68421	Aah68421 C glutami
C 7	822	34.6	822	4 AAF71779	Aaf71779 Coryneb
C 8	822	34.6	822	4 AAS96098	Aas96098 C. glutam
C 9	711	29.9	711	8 ACC80941	Acc80941 LysE prot
C 10	708	29.8	708	5 AAH68420	Aah68420 C glutami
C 11	698.6	29.4	1568	4 AAH45375	Aah45375 C. thermo
C 12	696.8	29.4	712	8 ACC80942	Acc80942 LysE24 pr
C 13	627	26.4	627	5 AAH68419	Aah68419 C glutami
C 14	485.2	20.4	993	4 AAF68077	Aaf68077 Coryneb
C 15	305.8	12.9	1035	7 ACA29651	ACA29651 Prokaryot
C 16	302.8	12.8	879	7 ACA29653	ACA29653 Prokaryot
C 17	132.6	5.6	1041	7 ACA25567	ACA25567 Prokaryot
C 18	123.8	5.2	897	7 ACA26879	ACA26879 Prokaryot
C 19	122.8	5.2	909	7 ACA38559	ACA38559 Prokaryot
C 20	122.8	5.2	15239	2 AAT33536	Aat33536 BCG delet
C 21	121.2	5.1	912	7 ACA40574	ACA40574 Prokaryot
C 22	121.2	5.1	110000	4 AAT99682_22	Continuation (23 o
C 23	121.2	5.1	110000	4 AAT99683_22	Continuation (23 o

24	116.4	4.9	1041	7	ACA24017	Prokaryot
25	114.8	4.8	1107	7	ACA26520	Prokaryot
26	108.6	4.6	1038	7	ACA32095	Prokaryot
C 27	108	4.5	1207	9	ADD13356	C. glutam
28	107	4.5	1041	7	ACA19150	Prokaryot
29	105.2	4.4	999	7	ACA51519	Prokaryot
30	105.2	4.4	999	7	ACA50894	Prokaryot
C 31	101	4.3	1200	4	AAF71729	Coryneb
C 32	100.8	4.2	5541	2	AAQ55755	Escherich
33	97.4	4.1	1035	7	ACA45509	Prokaryot
34	93.2	3.9	988	6	ABK72786	Bacillus
35	90.4	3.8	349980	6	ABQ81844	Bifidobac
36	82.6	3.5	894	4	AA525558	E. coli D
37	82.6	3.5	894	7	ACA32622	Prokaryot
38	81.4	3.4	999	7	ACA24884	Prokaryot
39	79.8	3.4	909	7	ACA53702	Prokaryot
40	79.4	3.3	23128	4	AA595552	Propionib
41	79.4	3.3	23128	7	ACF64481	Propionib
42	78.2	3.3	990	7	ACA35500	Prokaryot
C 43	78	3.3	1077	5	AAH66360	C glutami
44	77.6	3.3	903	7	ACF71729	Photobab
C 45	77.6	3.3	110000	7	ACF67367_49	Continuation (50 o

ALIGNMENTS

RESULT 1  
AAT96816/c  
ID AAT96816 standard; DNA; 2374 BP.  
XX  
AC AAT96816;  
XX  
DT 12-MAR-1998 (first entry)  
XX  
DE DNA encoding LysG, LysE and ORF3 from Corynebacterium glutamicum.  
XX  
KW LysG; LysE; ORF3; lysine transport; regulatory protein; export protein;  
Microbial production; amino acid; animal feed additive; ds.  
XX  
OS Corynebacterium glutamicum.  
XX  
FH Key Location/Qualifiers  
FT CDS complement(82..954)  
FT /tag= a  
FT /label= LysG  
FT 1016..1726  
FT /tag= b  
FT /label= LysE  
FT CDS complement(1723..2373)  
FT /tag= c  
FT /label= orf3  
DEI9548222-A1.  
26-JUN-1997.  
22-DEC-1995; 95DE-01048222.  
22-DEC-1995; 95DE-01048222.  
(KERJ ) FORSCHUNGSZENTRUM JUELICH GMBH.  
Vrljic M, Eggeling L, Sahm H;  
WPI; 1997-333867/31.  
P-PSDB; AAW37714, AAW37715, AAW37716.  
Increasing microbial production of amino acids, especially lysine - by  
improving export carrier activity or corresponding gene expression, also  
new export and regulatory genes from Corynebacterium.  
Claim 23 and 26; Page; 16pp; German.

XX CC This DNA, isolated from *Corynebacterium glutamicum*, contains the *lysG*,  
CC *lysE* and *ORF3* genes. *lysG* and *lysE* encode a lysine transport regulatory  
CC protein and an export protein, respectively. Microbial production of  
CC amino acids (A) is improved by increasing the export-carrier activity  
CC and/or the export gene expression in a microorganism that produces (A).  
CC The method is specifically used to increase production of lysine, used as  
CC an animal feed additive. Other (A) are variously useful as  
CC pharmaceuticals, condiments and intermediates for fine chemicals. This  
CC method increases the amount of (A) secreted into the culture medium.  
CC Export of (A) has been found to depend on a single gene. NB. This  
CC sequence has been created from the information given in table 2 of the  
CC specification  
XX CC  
SQ Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 2374; DB 2; Length 2374;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 AGATACCTCTTTGGAAGAACCATGTAGCAATTGGCGTACATTGTTGGCTCTGGAAGGC 60  
DB 2374 AGATCTCTCTTTGGAAGAACCATGTAGCAATTGGCGTACATTGTTGGCTCTGGAAGGC 2315  
  
QY 61 TCTTTACGTGGGTATTTCTTCTACCGTCCAGAGCTCACAGCGGAGGCGGTGAGTTTCA 120  
DB 2314 TCTTTAGGTGGGTATTTCTTCTACCGTCCAGAGCTCACAGCGGAGGCGGTGAGTTTCA 2255  
  
QY 121 GCGGAGGAGGGCTGCCCGCTTCTGATTCATCAGCCAGCTATTTCATCATTAATCGTTG 180  
DB 2254 GCGGAGGAGGGCTGCCCGCTTCTGATTCATCAGCCAGCTATTTCATCATTAATCGTTG 2195  
  
QY 181 GGTGAGGAGACCGGGCGATCAGCGTGAGAACTTGTGTCAGTCAGTCGCCAACAATGCT 240  
DB 2194 GGTGAGGAGACCGGGCGATCAGCGTGAGAACTTGTGTCAGTCAGTCGCCAACAATGCT 2135  
  
QY 241 TGGCGTCATTTGCTTCTCACCACTTGGCGAGGGCTGCTCACGGAACAATATCTCGATGG 300  
DB 2134 TGGCGTCATTTGCTTCTCACCACTTGGCGAGGGCTGCTCACGGAACAATATCTCGATGG 2075  
  
QY 301 AATTCCAGAGGGTCCCGCGCCAGCCAGCGGTAAGTCCCTGTCTGAGGGCATGTTGAAGT 360  
DB 2074 AATTCCAGAGGGTCCCGCGCCAGCCAGCGGTAAGTCCCTGTCTGAGGGCATGTTGAAGT 2015  
  
QY 361 GAACAATATTGATATGTCGCAAGCTCAATGACATCGCCAGGAACCGCGGCAGTCAC 420  
DB 2014 GAACAATATTGATATGTCGCAAGCTCAATGACATCGCCAGGAACCGCGGCAGTCAC 1955  
  
QY 421 TGGCGAGATGGCGCTTGATCGATGGGTGTCGCGAGCAAGAGAGATACGCGCGGATACCGT 480  
DB 1954 TGGCGAGATGGCGCTTGATCGATGGGTGTCGCGAGCAAGAGAGATACGCGCGGATACCGT 1895  
  
QY 481 GACCACTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
DB 1894 GACCACTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1835  
  
QY 541 CAACAACATTGGAGTTTCTGACGCGGAGTTGGAGGCGATCGATGAGATTTCCACGAGCG 600  
DB 1834 CAACAACATTGGAGTTTCTGACGCGGAGTTGGAGGCGATCGATGAGATTTCCACGAGCG 1775  
  
QY 601 CGGCATCAACATTTGGGCGAAGCGCACCGATTCGAAACCGCGGAAAACCTAACCCATCAA 660  
DB 1774 CGGCATCAACATTTGGGCGAAGCGCACCGATTCGAAACCGCGGAAAACCTAACCCATCAA 1715  
  
QY 661 CATCAGTTTGTAGGCGCAATGGGTGATCACTCACTGCACTGCCAGCAGCTGTGATCCAGCGCA 720  
DB 1714 CATCAGTTTGTAGGCGCAATGGGTGATCACTCACTGCACTGCCAGCAGCTGTGATCCAGCGCA 1655  
  
QY 721 CACCTTGGGCGTGACAGCGGGGTGACAAATGCTGCTGCCCGGAAACCCACGAGCGGAA 780  
DB 1654 CACCTTGGGCGTGACAGCGGGGTGACAAATGCTGCTGCCCGGAAACCCACGAGCGGAA 1595  
  
QY 781 CCAGATCAGGCTTGGCCGGAACGCGCCAGCGGCGGAAATCCACCGTCCGGTGTGCGCGGTA 840

DB 1594 CCAGATCAGGCTTGGCGGAGCGGCGAGCGGCGGAAATCCACCGCTCGGTGTGCGCGTA 1535  
QY 841 TTGGGCGCGAGCGCGCGGATATAACAAAACGCGTCCAAATAGCATTGGGTTCAACCA 900  
DB 1534 TTGGGCGCGAGCGCGCGGATATAACAAAACGCGTCCAAATAGCATTGGGTTCAACCA 1475  
QY 901 GGTTCAGCAGATTGTCATCAATCAATGGGCTTTACCCAAACCCGCTCTTATCGAGGCTCAC 960  
DB 1474 GGTTCAGCAGATTGTCATCAATCAATGGGCTTTACCCAAACCCGCTCTTATCGAGGCTCAC 1415  
QY 961 CTCACCCGCGAGCGGTGCGGTGTCAGTGGCCACCGCGCGAACCGCCAAAGCGGTGTC 1020  
DB 1414 CTCACCCGCGAGCGGTGCGGTGTCAGTGGCCACCGCGCGAACCGCCAAAGCGGTGTC 1355  
QY 1021 ATGGGCGAGGTTGCTGCTCTTCAATGATCTGTGGCGCTTCCACCTGTTGTGTCAT 1080  
DB 1354 ATGGGCGAGCGGTGCTGCTCTTCAATGATCTGTGGCGCTTCCACCTGTTGTGTCAT 1295  
QY 1081 GGGGTCTTTTGGTCCATGACGCGCAACCAATACAGGTAAAGCATGTCACCCAGCGCAT 1140  
DB 1294 GGGGTCTTTTGGTCCATGACGCGCAACCAATACAGGTAAAGCATGTCACCCAGCGCAT 1235  
QY 1141 AATATGACGACGATCGCGCGGCAATTGGACAAAAGATCAACGCGAAGGTGCGCGCAT 1200  
DB 1234 AATATGACGACGATCGCGCGGCAATTGGACAAAAGATCAACGCGAAGGTGCGCGCAT 1175  
QY 1201 GAACAAAAGAGCGTCAGAAATTAACACACGAGAAAGAACCGCAATGAGTCTTTCGCGCTT 1260  
DB 1174 GAACAAAAGAGCGTCAGAAATTAACACACGAGAAAGAACCGCAATGAGTCTTTCGCGCTT 1115  
QY 1261 AATTCTCTTTTAAATCACAGTACATTTGCGGTCGATGGAAGATGAGTAAAGATGTCGCGCC 1320  
DB 1114 AATTCTCTTTTAAATCACAGTACATTTGCGGTCGATGGAAGATGAGTAAAGATGTCGCGCC 1055  
QY 1321 CAAAGGACGCTGTATGAGATTTCCATGATCAGCATGCTGATGAGTAAAGATGAGTAAAGATGAGTAA 1380  
DB 1054 CAAAGGACGCTGTATGAGATTTCCATGATCAGCATGCTGATGAGTAAAGATGAGTAAAGATGAGTAA 995  
QY 1381 AGTAAATGATTTGTTTAAATGATGTTTAAATGATGTTTAAATGATGTTTAAATGATGTTTAAATGATGTT 1440  
DB 994 AGTAAATGATTTGTTTAAATGATGTTTAAATGATGTTTAAATGATGTTTAAATGATGTTTAAATGATGTT 935  
QY 1441 CACTTTGCTCTCAATCATGATGAGAGCGCTTGAAGGCGCTTCTTAGCCCTTTCCAT 1500  
DB 934 CACTTTGCTCTCAATCATGATGAGAGCGCTTGAAGGCGCTTCTTAGCCCTTTCCAT 875  
QY 1501 TTCCCTCTCGCGGTGATGAGCGGTTAAAGCTCTGAGCATCACTGAGTGGGTGAGTGT 1560  
DB 874 TTCCCTCTCGCGGTGATGAGCGGTTAAAGCTCTGAGCATCACTGAGTGGGTGAGTGT 815  
QY 1561 GGTATCGCGCACCCAAACCGGCGAAAGCAACCGAAGCGGGTGAAGTCTTGTGCAAGCAGC 1620  
DB 814 GGTATCGCGCACCCAAACCGGCGAAAGCAACCGAAGCGGGTGAAGTCTTGTGCAAGCAGC 755  
QY 1621 GCGGAAATGTTGTTGTCGACGAGAACTTAAGCGCAACTATCTGAGCGCTTGTGGA 1680  
DB 754 GCGGAAATGTTGTTGTCGACGAGAACTTAAGCGCAACTATCTGAGCGCTTGTGGA 695  
QY 1681 AATCCGTTAAACCATCGCATCAACGAGATTCGCTATCCCATGTTTCTCCCGTGT 1740  
DB 694 AATCCGTTAAACCATCGCATCAACGAGATTCGCTATCCCATGTTTCTCCCGTGT 635  
QY 1741 CAAGAGTAGTCTTCTGGGTGGAGCAACGCTCACGCTCGCTTGGAGATGAAGCGCA 1800  
DB 634 CAAGAGTAGTCTTCTGGGTGGAGCAACGCTCACGCTCGCTTGGAGATGAAGCGCA 575  
QY 1801 CACATTTATCTTGTGCGGCGTGAAGTGTGTTAGGAGCGGTAACCGTGAAGTATTC 1860  
DB 574 CACATTTATCTTGTGCGGCGTGAAGTGTGTTAGGAGCGGTAACCGTGAAGTATTC 515  
QY 1861 GTCGCGGATGTAAGTAGTAGTCTTGAAACCATGCGCATCTTGGCGATTTGCAACCC 1920

Db 514 CCGTGGCGGATGTGAAGTAGTAGAACTTGGAAACCATGCGCCACCTTGGCCATTTGGAACCCC 455  
 QY 1921 CTCATTGGGGGATGCTACATGCTGTGATGGAACTAGATTGGGCTGGCATCCCGCTTT 1980  
 Db 454 CTCATTGGGGGATGCTACATGCTGTGATGGAACTAGATTGGGCTGGCATCCCGCTTT 395  
 QY 1981 AGCTTTCGGTCCCAAGATGTCTTCAAGACCGTGACCTGGACGGGCGGTGATGGTCC 2040  
 Db 394 AGCTTTCGGTCCCAAGATGTCTTCAAGACCGTGACCTGGACGGGCGGTGATGGTCC 335  
 QY 2041 TGTGGGCGCAGCGGCTATCATGCTCCCGTCCGCGAAGTTTGTGTGGCAATTCG 2100  
 Db 334 TGTGGGCGCAGCGGCTATCATGCTCCCGTCCGCGAAGTTTGTGTGGCAATTCG 275  
 QY 2101 CCGAGGCTTGTGGGCACTTCTCCGGAACCCCAAGCTGCTCCCATCTAAGACAGG 2160  
 Db 274 CCGAGGCTTGTGGGCACTTCTCCGGAACCCCAAGCTGCTCCCATCTAAGACAGG 215  
 QY 2161 AGAAGTGATCTCTCGATGAGATACCATTTGACACACCGATGTATTGGCAACGATGGG 2220  
 Db 214 AGAAGTGATCTCTCGATGAGATACCATTTGACACACCGATGTATTGGCAACGATGGG 155  
 QY 2221 CCTGGATCTAGATCTCTAGCTAGATCTACAGACCGCTGTGATGAGCAATTCGAGG 2280  
 Db 154 CCTGGATCTAGATCTCTAGCTAGATCTACAGACCGCTGTGATGAGCAATTCGAGG 95  
 QY 2281 ATTGGGCGCTTACTTCTGAAAGGTTTCAGGGTTTTTCACTTCTGCGCCGCGAGAA 2340  
 Db 94 ATTGGGCGCTTACTTCTGAAAGGTTTCAGGGTTTTTCACTTCTGCGCCGCGAGAA 35  
 QY 2341 TTGGGCCAGCAGAGTAACACCTTCAGCAAAATGG 2374  
 Db 34 TTGGGCCAGCAGAGTAACACCTTCAGCAAAATGG 1

RESULT 2  
 ADB66196/c  
 ID ADB66196 standard; DNA; 2374 BP.  
 XX  
 AC ADB66196;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE DNA fragment containing C. glutamicum LysE and LysE genes.  
 XX  
 KW L-arginine production; coryneform bacteria; LysE; arginine repressor;  
 KW argR; liver function promoting agent; amino acid infusion;  
 KW amino acid pharmaceutical; LysE; ds.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 FH Key Location/Qualifiers  
 XX CDS 1025..1726  
 FT /\*tag= a  
 FT /product= "protein encoded by LysE gene"  
 XX  
 FN US2003113899-A1.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 17-JUL-2002; 2002US-00196232.  
 XX  
 PR 25-JUL-2001; 2001JP-00224586.  
 XX  
 PA (AJIN ) AJINOMOTO CO INC.  
 XX  
 PI Yamauchi M, Ito H, Gunji Y, Yasueda H;  
 XX  
 DR WPI; 2003-708853/67.  
 DR P-PSDB; ADB66197.  
 XX  
 PT A microorganism comprising enhanced expression of the lysE gene is useful  
 for enhanced production of L-arginine.

XX Example 4; Page 26-27; 36pp; English.  
 PS  
 CC The present invention relates to a method for producing L-arginine in a  
 CC microorganism (e.g. coryneform bacteria) that has L-arginine producing  
 CC ability and has been modified for enhanced expression of the lysE gene.  
 CC The microorganism is also modified so that an arginine repressor (argR)  
 CC does not function normally. The method of the invention is useful for the  
 CC enhanced production of L-arginine which is useful in liver function  
 CC promoting agents, amino acid infusion and comprehensive amino acid  
 CC pharmaceuticals. The present sequence represents a DNA fragment  
 CC containing Corynebacterium glutamicum LysG and LysE genes. Note: The  
 CC present sequence is given as SEQ ID No.24 in the Sequence Listing but is  
 CC referred to as SEQ ID No.25 in the rest of the specification.  
 XX  
 SQ Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 2374; DB 9; Length 2374;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGAATCTCTTTGGAAGAACCACTGTAGCATTTGCGTACATTTTTCGCTTCGGAAGGC 60  
 Db 2374 AGATACTCTTTGGAAGAACCACTGTAGCATTTGCGTACATTTTTCGCTTCGGAAGGC 2315  
 QY 61 TCTTTACGTGGGTATTTCTTCTACGGTCCAGAGCTTCACAGCGAGCGGCTGAGTTTCA 120  
 Db 2314 TCTTTACGTGGGTATTTCTTCTACGGTCCAGAGCTTCACAGCGAGCGGCTGAGTTTCA 2255  
 QY 121 GCGGAGGAGGGTTCGCCGCTTCTGATTCATCAGCAAGCTATTCCATCAATTAATCGTTG 180  
 Db 2254 GCGGAGGAGGGTTCGCCGCTTCTGATTCATCAGCAAGCTATTCCATCAATTAATCGTTG 2195  
 QY 181 GGTGAGGAACCGGGCGATGACGGTGTGAACTTTGTGAGTCAGTCAGTCGCAACATGGTCT 240  
 Db 2194 GGTGAGGAACCGGGCGATGACGGTGTGAACTTTGTGAGTCAGTCAGTCGCAACATGGTCT 2135  
 QY 241 TGGCGTCAATTGCTTCTCACCACTTGGCGAGGGCTGTCTACGCAACAAATATCTCGATGG 300  
 Db 2134 TGGCGTCAATTGCTTCTCACCACTTGGCGAGGGCTGTCTACGCAACAAATATCTCGATGG 2075  
 QY 301 AATTCAGAGGGTTCGCCGCGCAGCAGGGTAAAGTCCCTGTCTGAGGGCAATGTTGAAAGT 360  
 Db 2074 AATTCAGAGGGTTCGCCGCGCAGCAGGGTAAAGTCCCTGTCTGAGGGCAATGTTGAAAGT 2015  
 QY 361 GACCAATTTGATATGGTTCGGCAGCTCATGACATCCGCCAGCAACGCGGCGCTCACT 420  
 Db 2014 GACCAATTTGATATGGTTCGGCAGCTCATGACATCCGCCAGCAACGCGGCGCTCACT 1955  
 QY 421 TGGCGAGATGGCGCTTGCATGGGTGCTGCGCAGCAAGAGAGTACGCGCGGATACCGT 480  
 Db 1954 TGGCGAGATGGCGCTTGCATGGGTGCTGCGCAGCAAGAGAGTACGCGCGGATACCGT 1895  
 QY 481 GACCAATTTGATATGGTTCGGCAGCTCATGACATCCGCCAGCAACGCGCTTCACT 540  
 Db 1894 GACCAATTTGATATGGTTCGGCAGCTCATGACATCCGCCAGCAACGCGCTTCACT 1835  
 QY 541 CAACAATTTGATATGGTTCGGCAGCTCATGACATCCGCCAGCAACGCGCTTCACT 600  
 Db 1834 CAACAATTTGATATGGTTCGGCAGCTCATGACATCCGCCAGCAACGCGCTTCACT 1775  
 QY 601 CCGCATCAACATTTGGCGAAGGCGCACCGATTCAAAAACCGCGAATACTAACCCATCAA 660  
 Db 1774 CCGCATCAACATTTGGCGAAGGCGCACCGATTCAAAAACCGCGAATACTAACCCATCAA 1715  
 QY 661 CATCATTTGATATGGGCAATGGGTTCATCACTGCGCAGCAGCTTGTATCCAGCGCA 720  
 Db 1714 CATCATTTGATATGGGCAATGGGTTCATCACTGCGCAGCAGCTTGTATCCAGCGCA 1655  
 QY 721 CACCTTGGGGCTGGACAGCGGGGTGACATGCTGCGCCGAGAAACCCACAGCGGAA 780  
 Db 1654 CACCTTGGGGCTGGACAGCGGGGTGACATGCTGCGCCGAGAAACCCACAGCGGAA 1595

QY	781	CCAGATCAGGCTTGC	CGCGAAACGCGCCAGCGCGGAGAAATCCACCGTCCGGTGFPGCGCGTA	840
Db	1594	CCAGATCAGGCTTGC	CGCGGAAACGCGCGCACGCGCGGAAATCCACCGTCCGGTGFPGCGCGTA	1535
QY	841	TTGCGCGCGAGCGCGCGGATAAACACAAACGCGTCCAAATACGATTTCCGGTTCACACCA	900	
Db	1534	TTGCGCGCGAGCGCGCGGATAAACACAAACGCGTCCAAATACGATTTCCGGTTCACACCA	1475	
QY	901	GGTCAGCACGATTTGC	ATCATCAACATPAGGGCTTTACCCAAACCGCGTCTTATCGACGCTCAC	960
Db	1474	GGTCAGCACGATTTGC	ATCATCAACATPAGGGCTTTACCCAAACCGCGTCTTATCGACGCTCAC	1415
QY	961	CTCCACCGCGACCGGTTG	CGCGTGTCAGTGGSCACCGCGACCGCGACCGCGCCAAAGGCGTGC	1020
Db	1414	CTCCACCGCGACCGGTTG	CGCGTGTCAGTGGSCACCGCGACCGCGACCGCGCCAAAGGCGTGC	1355
QY	1021	ATCGGCGACGTTGGTTC	TCTGTCTTCCAAATGATCTGTGGCGCTTCCACCTGTGTTGTGCAT	1080
Db	1354	ATCGGCGACGTTGGTTC	TCTGTCTTCCAAATGATCTGTGGCGCTTCCACCTGTGTTGTGCAT	1295
QY	1081	GGCGTCTTTGCGTGC	ATGATGACGGCAACCATAAACGATTAAGCGATGCGACCGCAGCGCAT	1140
Db	1294	GGCGTCTTTGCGTGC	ATGATGACGGCAACCATAAACGATTAAGCGATGCGACCGCAGCGCAT	1235
QY	1141	AAATATCGAGCACGATCGG	CGCGGCATTTGGAACAAAGATCAAACGCCCAAGGTGCGCGCGCAT	1200
Db	1234	AAATATCGAGCACGATCGG	CGCGGCATTTGGAACAAAGATCAAACGCCCAAGGTGCGCGCGCAT	1175
QY	1201	GAAACAAAGACGTGAGA	AAATTAACACACGAGAGAAACCGCATAGTCTTCCTCGCGCTT	1260
Db	1174	GAAACAAAGACGTGAGA	AAATTAACACACGAGAGAAACCGCATAGTCTTCCTCGCGCTT	1115
QY	1261	AAATTCCTTTGTTTAAT	CACACAGTACATTTCTGCGGTCGATGGACAGTAAGAAAGATGCGGCC	1320
Db	1114	AAATTCCTTTGTTTAAT	CACACAGTACATTTCTGCGGTCGATGGACAGTAAGAAAGATGCGGCC	1055
QY	1321	CAAAAGCAGACGTGTAAT	GAGATTTCCATGATCACCATCTGTGACCTATGGAAGTACTTA	1380
Db	1054	CAAAAGCAGACGTGTAAT	GAGATTTCCATGATCACCATCTGTGACCTATGGAAGTACTTA	995
QY	1381	AGTAAATATGATGGTTC	TTTAAATAGCTTCAATAGCTTCATGAACCCCATTCACATCGSA	1440
Db	994	AGTAAATATGATGGTTC	TTTAAATAGCTTCAATAGCTTCATGAACCCCATTCACATCGSA	935
QY	1441	CATTTTGCTCTCAATCA	TATGATGAAGCAGCTTCGAAGCGCCTCCTTTAGCCCTTTCCAT	1500
Db	934	CATTTTGCTCTCAATCA	TATGATGAAGCAGCTTCGAAGCGCCTCCTTTAGCCCTTTCCAT	875
QY	1501	TTTCCCGCTCGCGCGT	CAGTACGCGTTAAAGCTCTCGACATCAAGTGAGGTGAGTGT	1560
Db	874	TTTCCCGCTCGCGCGT	CAGTACGCGTTAAAGCTCTCGACATCAAGTGAGGTGAGTGT	815
QY	1561	GGTATCGCGCACCAAC	CGGCGCAAGCAACCGAGCGGTTGAAGTCTTTGTGCAAGCAGC	1620
Db	814	GGTATCGCGCACCAAC	CGGCGCAAGCAACCGAGCGGTTGAAGTCTTTGTGCAAGCAGC	755
QY	1621	GCGGAAATAGTGTCTG	CGAAGCAGAACTAAAGCGCAATATCTGAGCGCTTCTGTA	1680
Db	754	GCGGAAATAGTGTCTG	CGAAGCAGAACTAAAGCGCAATATCTGAGCGCTTCTGTA	695
QY	1681	AAATCCCGTTAACCAT	CGCCATCGCATCAACGAGATTCGCTATCCACATGTTTCTCCCGTGT	1740
Db	694	AAATCCCGTTAACCAT	CGCCATCGCATCAACGAGATTCGCTATCCACATGTTTCTCCCGTGT	635
QY	1741	CAACAGGATAGCTTCT	TGCGGTFGAGCAACGCTCAACGCTCGCTTTGGAAGATGAAGCGCA	1800
Db	634	CAACAGGATAGCTTCT	TGCGGTFGAGCAACGCTCAACGCTCGCTTTGGAAGATGAAGCGCA	575
QY	1801	CACATATTCCTTTCG	CGCGGTGGAGATTTTTAGAGCGGTAAACCGGTGAAGCTAATCC	1860
Db	574	CACATATTCCTTTCG	CGCGGTGGAGATTTTTAGAGCGGTAAACCGGTGAAGCTAATCC	515
QY	1861	CGTGGCGGGATGTGA	AGTAAACTTGGAAACCACTTCGCGCACTTGGCGCATTTGGCAACCC	1920

Db	514	CGTGGCGGGATGTGAAGTAGTAGAATCTTGGAAACATCGGCCACTTTGGCGCAATTGGCAACCCC	455
Qy	1921	CTCATTTGGCGGATGCTCATCATGTTGATGGGAACTAGATTGGGCTGCGATGCCGCTTT	1980
Db	454	CTCATTTGGCGGATGCTCATCATGTTGATGGGAACTAGATTGGGCTGCGATGCCGCTTT	395
Qy	1981	AGCCTTTGGTCCCAAGATGTGCTTCAAGACCGTGACCTTGGACGGGCGCTCGATGTC	2040
Db	394	AGCCTTTGGTCCCAAGATGTGCTTCAAGACCGTGACCTTGGACGGGCGCTCGATGTC	335
Qy	2041	TGTGGGGCGACGCGCGTATCCATTGTCCTGGCGGAGAGTTTGGTGGAGCAATTG	2100
Db	334	TGTGGGGCGACGCGCGTATCCATTGTCCTGGCGGAGAGTTTGGTGGAGCAATTG	275
Qy	2101	CGAGGCGCTTGGTTGGGAGCTTCTCCGAAACCAAGCTGCTCCATGCTAAAGCAGG	2160
Db	274	CGAGGCGCTTGGTTGGGAGCTTCTCCGAAACCAAGCTGCTCCATGCTAAAGCAGG	215
Qy	2161	AGAATGATCTCTCGATGAGATACCCATTGACACACCGATGTTATGGCAAGATGCG	2220
Db	214	AGAATGATCTCTCGATGAGATACCCATTGACACACCGATGTTATGGCAAGATGCG	155
Qy	2221	CTTGGAAATCTAGATCTCTAGCTAGACTCAAGACCGCGCTGTTGATGAGCAATCGAGG	2280
Db	154	CTTGGAAATCTAGATCTCTAGCTAGACTCAAGACCGCGCTGTTGATGAGCAATCGAGG	95
Qy	2281	ATTGGGGCGCTTAGTACTCTGTAAAGGTTGAGGTTTTTCACTTCTTCGCCCGCAGAA	2340
Db	94	ATTGGGGCGCTTAGTACTCTGTAAAGGTTGAGGTTTTTCACTTCTTCGCCCGCAGAA	35
Qy	2341	TTGGGCCAGGCAGGTAAACCTTTGAGCAATGG	2374
Db	34	TTGGGCCAGGCAGGTAAACCTTTGAGCAATGG	1

RESULT 3	
AAH68528	
ID	AAH68528 standard; DNA; 349980 BP.
XX	
AC	AAH68528;
XX	
DT	26-SEP-2001 (first entry)
XX	
DE	C glutamicum coding sequence fragment SEQ ID NO: 7063.
XX	
KW	Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX	
KW	organic acid synthesis; ds.
XX	
OS	Corynebacterium glutamicum.
XX	
PN	EP1108790-A2.
XX	
PD	20-JUN-2001.
XX	
PF	18-DEC-2000; 2000EP-00127688.
XX	
PR	16-DEC-1999; 99JP-00377484.
PR	07-APR-2000; 2000JP-00159162.
PR	03-AUG-2000; 2000JP-00280988.
XX	
PA	(KYOW ) KYOWA HAKKO KOGYO KK.
XX	
PI	Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI	Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX	
DR	WPI; 2001-376931/40.
XX	
PT	Novel polynucleotides derived from Corynebacterium, for identifying
PT	mutation point of a gene, measuring expression of a gene, analyzing
PT	expression profile or pattern of a gene and identifying homologous gene.
XX	
PS	Disclosure; SEQ ID NO 7063; 246pp + Sequence Listing; English.

xx The present invention provides a number of nucleotide and protein  
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of corynebacterium bacterium, measuring expression amount and analysing  
 CC the expression profile or expression pattern of a gene derived from  
 CC corynebacterium bacterium, and identifying a homologue of a gene derived from  
 CC corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino  
 CC acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a nucleic acid described  
 CC in the exemplification of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from the European Patent Office  
 xx  
 xx Sequence 349980 BP; 80727 A; 91049 C; 97618 G; 80586 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 2374; DB 5; Length 349980;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGATACCTCTTTCGAGAGAACCATGTACGATTCGCTGACATTTGTTGGCTCTCGAAGGC 60  
 Db 127595 AGATACCTCTTTCGAGAGAACCATGTACGATTCGCTGACATTTGTTGGCTCTCGAAGGC 127654  
 QY 61 TCTTTACGTGGGTATTTCTTCTTACGGTCCAGAGCTCACAGCGAGGCGGCTGAGTTTAT 120  
 Db 127655 TCTTTACGTGGGTATTTCTTCTTACGGTCCAGAGCTCACAGCGAGGCGGCTGAGTTTAT 127714  
 QY 121 GCGCGAGAGGGTCCCGCTTCTGATTCATCATGAGCAAGCTATTCATCATTAATTCGTTG 180  
 Db 127715 GCGCGAGAGGGTCCCGCTTCTGATTCATCATGAGCAAGCTATTCATCATTAATTCGTTG 127774  
 QY 181 GGTGAGAGAACCGGCGATGACGCTGAGAACTTTGTCAGTCAAGCTGCCCAACATGTCT 240  
 Db 127775 GGTGAGAGAACCGGCGATGACGCTGAGAACTTTGTCAGTCAAGCTGCCCAACATGTCT 127834  
 QY 241 TGGCGTCAATGCTTTTCCACCACTTGCAGAGGCGCTGCTCACGACAAATATCTCGATGG 300  
 Db 127835 TGGCGTCAATGCTTTTCCACCACTTGCAGAGGCGCTGCTCACGACAAATATCTCGATGG 127894  
 QY 301 AATTCCAGAGGTTCCCGGCGACGAGGTAAGTCCCTGTCTGAGGCGATGTTGAACGT 360  
 Db 127895 AATTCCAGAGGTTCCCGGCGACGAGGTAAGTCCCTGTCTGAGGCGATGTTGAACGT 127954  
 QY 361 GAACAATATTGATATGCTCCGCAAGCTCAATGACATCCGCCAGAACCGGCGAGTCACT 420  
 Db 127955 GAACAATATTGATATGCTCCGCAAGCTCAATGACATCCGCCAGAACCGGCGAGTCACT 128014  
 QY 421 TGGCGAGATGGGCTTGATGGGTGCTGCGGAGCAAGAGAGTACGCGCGGATACCGT 480  
 Db 128015 TGGCGAGATGGGCTTGATGGGTGCTGCGGAGCAAGAGAGTACGCGCGGATACCGT 128074  
 QY 481 GACCAATGATGATTTGTTGCTTCTGCTGAGTCAAGCTGAGCAACAGCCTTGATTCAT 540  
 Db 128075 GACCAATGATGATTTGTTGCTTCTGCTGAGTCAAGCTGAGCAACAGCCTTGATTCAT 128134  
 QY 541 CAACAACCTTGGAGTTTCTGAGCGGAGTTGGAGGCGATCGATGAGATTTCCACGAGC 600  
 Db 128135 CAACAACCTTGGAGTTTCTGAGCGGAGTTGGAGGCGATCGATGAGATTTCCACGAGC 128194  
 QY 601 CGGATCAACATTTGGGCGAAGGCCACCGATTCGAAACCGGCGAATACTAACCCATCAA 660  
 Db 128195 CGGATCAACATTTGGGCGAAGGCCACCGATTCGAAACCGGCGAATACTAACCCATCAA 128254  
 QY 661 CATCAGTTTCAATGCGGATGCGTCACTCACTCCAGCAGCAGCTTGATCCAGGCCA 720  
 Db 128255 CATCAGTTTCAATGCGGATGCGGATGCGTCACTCACTCCAGCAGCAGCTTGATCCAGGCCA 128314  
 QY 721 CACTTGGGCTGGACAGCGGCGTGAATGCTGCTGCGCGGAAACCCACCGAGCGGAA 780  
 Db 128315 CACTTGGGCTGGACAGCGGCGTGAATGCTGCTGCGCGGAAACCCACCGAGCGGAA 128374  
 QY 781 CCAGATCAGGCTTGGCGGAACCGCGCAGCGGAGAAATCCACCGTCCGGTGTGCGCGTA 840

Db 128375 CCAGATCAGGTTTGGCGGAGAACCGCCAGCGGAGAAATCCACGTCGCTGTGCGCGTA 128434  
 QY 841 TTGCGCGCGAGCGCGCGGATTAACACAAACCGCTCCAAATAAGCATTTGGGTTCAACCA 900  
 Db 128435 TTGCGCGCGAGCGCGCGGATTAACACAAACCGCTCCAAATAAGCATTTGGGTTCAACCA 128494  
 QY 901 GGTGAGCAAGTATGCGCATCAAGATGGGCTTACCGAACCCTGCTTATCGACGCTCAC 960  
 Db 128495 GGTGAGCAAGTATGCGCATCAAGATGGGCTTACCGAACCCTGCTTATCGACGCTCAC 128554  
 QY 961 CTCACCGCGACCGGCTTGGGCTGTCAGTGGCCACCGCGAAGCGCCAAAGGGGTGTC 1020  
 Db 128555 CTCACCGCGACCGGCTTGGGCTGTCAGTGGCCACCGCGAAGCGCCAAAGGGGTGTC 128614  
 QY 1021 ATCGGGACGCTGTTGTTCTGTTCTTCAATGATCTGTGGCGCTTCCACCTTGTGTTGTCAT 1080  
 Db 128615 ATCGGGACGCTGTTGTTCTGTTCTTCAATGATCTGTGGCGCTTCCACCTTGTGTTGTCAT 128674  
 QY 1081 GCGGCTTTTGGCTGCCATGACGGCAACCAATACAGGTAAGGATGCCACCCAGCGCAT 1140  
 Db 128675 GCGGCTTTTGGCTGCCATGACGGCAACCAATACAGGTAAGGATGCCACCCAGCGCAT 128734  
 QY 1141 AATATCGAGCACGATCGGCGCGCATTTGGAACAAAGATCAAGCGCCCAAGGTGCGCGCAT 1200  
 Db 128735 AATATCGAGCACGATCGGCGCGCATTTGGAACAAAGATCAAGCGCCCAAGGTGCGCGCAT 128794  
 QY 1201 GAACAAAAGACGTGCAAAATTAACACACAGAGAACCGGCAATGAGTCTTTCGCGCTT 1260  
 Db 128795 GAACAAAAGACGTGCAAAATTAACACACAGAGAACCGGCAATGAGTCTTTCGCGCTT 128854  
 QY 1261 AATTCCTTGTGTTAATCACCAGTACATTTCTCGGCTCGATGACAGTAAGAAAGACTGGCCCC 1320  
 Db 128855 AATTCCTTGTGTTAATCACCAGTACATTTCTCGGCTCGATGACAGTAAGAAAGACTGGCCCC 128914  
 QY 1321 CAAAAGCAGACCTGTAAATGAGATTTCCATGATCAACCTCGTGAAGTACTTTA 1380  
 Db 128915 CAAAAGCAGACCTGTAAATGAGATTTCCATGATCAACCTCGTGAAGTACTTTA 128974  
 QY 1381 AGTAAATGATGGTCTTAAATGAGATTTAAATAGTCTTCAAGACCCCATTCAGTCACTGA 1440  
 Db 128975 AGTAAATGATGGTCTTAAATGAGATTTAAATAGTCTTCAAGACCCCATTCAGTCACTGA 129034  
 QY 1441 CACTTTGCTCTCAATCATTTGATGAGGACGCTTCGAGGCGCTCTTAGCCCTTCCAT 1500  
 Db 129035 CACTTTGCTCTCAATCATTTGATGAGGACGCTTCGAGGCGCTCTTAGCCCTTCCAT 129094  
 QY 1501 TTTCCCTTGGCGGCTGAGTCAAGCGGTTAAAGCTCTCGAGCATCACTGTTGGTGGAGTGT 1560  
 Db 129095 TTTCCCTTGGCGGCTGAGTCAAGCGGTTAAAGCTCTCGAGCATCACTGTTGGTGGAGTGT 129154  
 QY 1561 GGTATCGCGACCCAAACCGGCGCAAGCAACCGAGGCGGTGAAGTCTTGTGCAAGCAGC 1620  
 Db 129155 GGTATCGCGACCCAAACCGGCGCAAGCAACCGAGGCGGTGAAGTCTTGTGCAAGCAGC 129214  
 QY 1621 GCGGAAATGCTGCTCAAGCAGAGAACTAAAGCGCAACTATCTGAGCGCTTCTGTA 1680  
 Db 129215 GCGGAAATGCTGCTCAAGCAGAGAACTAAAGCGCAACTATCTGAGCGCTTCTGTA 129274  
 QY 1681 AATCCGTTAAACCATCGCCATCAACGAGATTCGCTATCCATGTTTCTTCCCGTGT 1740  
 Db 129275 AATCCGTTAAACCATCGCCATCAACGAGATTCGCTATCCATGTTTCTTCCCGTGT 129334  
 QY 1741 CAAACGAGTACTCTTGGGCTGGAGCAACGCTCAAGTCTGAGTCTGAGAGTGAAGCGCA 1800  
 Db 129335 CAAACGAGTACTCTTGGGCTGGAGCAACGCTCAAGTCTGAGTCTGAGAGTGAAGCGCA 129394  
 QY 1801 CACATTATCTCTGCTGCGCGGTGAGATGTTTTAGGAGCGGTAAACCGGTGAAGTATCC 1860  
 Db 129395 CACATTATCTCTGCTGCGCGGTGAGATGTTTTAGGAGCGGTAAACCGGTGAAGTATCC 129454  
 QY 1861 CGTGGCGGATGTGAAGTAGTAGAACTTGGAAACCATGCGCACTTGGCCATTTGCAACCCC 1920

Db 129455 CGTGGCGGATGTGAAGTAGTAACTTGGAAACCATGCGCCACTTGGCCATGTCACACCC 129514  
 QY 1921 CTCATTGCGGATCCCTACATGTTGATGGAACTAGATTGGGCTCGATGCGCTTT 1980  
 Db 129515 CTCATTGCGGATCCCTACATGTTGATGGAACTAGATTGGGCTCGATGCGCTTT 129574  
 QY 1981 ACSCTTCGGTCCCAAGATGTCTTCAAGACCGTGCCTGGACGGGGCGCTCGATGGTCC 2040  
 Db 129575 ACSCTTCGGTCCCAAGATGTCTTCAAGACCGTGCCTGGACGGGGCGCTCGATGGTCC 129634  
 QY 2041 TGTGGGCGCGAGCGGCTATCCATTGTCCTGCGGCGAAGGTTTGGTGAGGCAATTCCG 2100  
 Db 129635 TGTGGGCGCGAGCGGCTATCCATTGTCCTGCGGCGAAGGTTTGGTGAGGCAATTCCG 129694  
 QY 2101 CCGAGGCTTGGTGGGGCTTCTCCGAAACCCAGCTCTCCCATGCTAAAGCAGG 2160  
 Db 129695 CCGAGGCTTGGTGGGGCTTCTCCGAAACCCAGCTCTCCCATGCTAAAGCAGG 129754  
 QY 2161 AGAAGTGATCTCTCTCGATGAGATACCCATTGACACCGATGATTTGGCAACGATGGCG 2220  
 Db 129755 AGAAGTGATCTCTCTCGATGAGATACCCATTGACACCGATGATTTGGCAACGATGGCG 129814  
 QY 2221 CTTGGAATCTAGATCTCTAGCTAGACTCAGAGCGGCTGTTGATGAGCAATGAGGG 2280  
 Db 129815 CTTGGAATCTAGATCTCTAGCTAGACTCAGAGCGGCTGTTGATGAGCAATGAGGG 129874  
 QY 2281 ATTGGCGGCTTAGTTACTTCTGAAAAGGTTTCAGGGTTTTCACCTTCTTCGCGCAGGAA 2340  
 Db 129875 ATTGGCGGCTTAGTTACTTCTGAAAAGGTTTCAGGGTTTTCACCTTCTTCGCGCAGGAA 129934  
 QY 2341 TTGGGCGGCGAGGATGAACCTTTCAGCAATGG 2374  
 Db 129935 TTGGGCGGCGAGGATGAACCTTTCAGCAATGG 129968

RESULT 4

AAAF17177  
 ID AAFA171777 standard; DNA; 993 BP.  
 XX  
 AC AAFA171777;  
 XX  
 DT 30-APR-2001 (first entry)  
 XX  
 DE Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:49.  
 XX  
 KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;  
 KW fine chemical production; microorganism; organic acid; nucleoside;  
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;  
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
 KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN WO200100843-A2.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 23-JUN-2000; 2000WO-IB000923.  
 XX  
 PR 25-JUN-1999; 99US-0141031P.  
 PR 01-JUL-1999; 99DE-01030476.  
 PR 02-JUL-1999; 99US-0142101P.  
 PR 08-JUL-1999; 99DE-01031415.  
 PR 08-JUL-1999; 99DE-01031418.  
 PR 08-JUL-1999; 99DE-01031419.  
 PR 08-JUL-1999; 99DE-01031420.  
 PR 08-JUL-1999; 99DE-01031424.  
 PR 08-JUL-1999; 99DE-01031428.  
 PR 08-JUL-1999; 99DE-01031434.  
 PR 08-JUL-1999; 99DE-01031435.  
 PR 08-JUL-1999; 99DE-01031443.  
 PR 08-JUL-1999; 99DE-01031453.  
 PR 08-JUL-1999; 99DE-01031457.

PR 08-JUL-1999; 99DE-01031465.  
 PR 08-JUL-1999; 99DE-01031478.  
 PR 08-JUL-1999; 99DE-01031510.  
 PR 08-JUL-1999; 99DE-01031541.  
 PR 08-JUL-1999; 99DE-01031573.  
 PR 08-JUL-1999; 99DE-01031592.  
 PR 08-JUL-1999; 99DE-01031632.  
 PR 08-JUL-1999; 99DE-01031634.  
 PR 08-JUL-1999; 99DE-01031636.  
 PR 08-JUL-1999; 99DE-01032125.  
 PR 09-JUL-1999; 99DE-01032126.  
 PR 09-JUL-1999; 99DE-01032130.  
 PR 09-JUL-1999; 99DE-01032186.  
 PR 09-JUL-1999; 99DE-01032206.  
 PR 09-JUL-1999; 99DE-01032227.  
 PR 09-JUL-1999; 99DE-01032228.  
 PR 09-JUL-1999; 99DE-01032229.  
 PR 09-JUL-1999; 99DE-01032230.  
 PR 14-JUL-1999; 99DE-01032922.  
 PR 14-JUL-1999; 99DE-01032926.  
 PR 14-JUL-1999; 99DE-01032928.  
 PR 14-JUL-1999; 99DE-01033004.  
 PR 14-JUL-1999; 99DE-01033005.  
 PR 14-JUL-1999; 99DE-01033006.  
 PR 27-AUG-1999; 99DE-01040764.  
 PR 27-AUG-1999; 99DE-01040765.  
 PR 27-AUG-1999; 99DE-01040766.  
 PR 27-AUG-1999; 99DE-01040832.  
 PR 31-AUG-1999; 99DE-01041378.  
 PR 31-AUG-1999; 99DE-01041379.  
 PR 31-AUG-1999; 99DE-01041380.  
 PR 31-AUG-1999; 99DE-01041394.  
 PR 31-AUG-1999; 99DE-01041396.  
 PR 03-SEP-1999; 99DE-01042076.  
 PR 03-SEP-1999; 99DE-01042077.  
 PR 03-SEP-1999; 99DE-01042079.  
 PR 03-SEP-1999; 99DE-01042086.  
 PR 03-SEP-1999; 99DE-01042087.  
 PR 03-SEP-1999; 99DE-01042088.  
 PR 03-SEP-1999; 99DE-01042095.  
 PR 03-SEP-1999; 99DE-01042124.  
 PR 03-SEP-1999; 99DE-01042129.  
 PR 09-MAR-2000; 2000US-0187970P.  
 XX  
 (BADI ) BASF AG.  
 PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
 PI WPI; 2001-137957/14.  
 XX P-PSDB; AAB79658.  
 DR  
 DR Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway  
 XX proteins, useful for producing fine chemicals in microorganisms,  
 XX including organic acids, nonproteinogenic amino acids, and purine and  
 XX pyrimidine bases.  
 PS Claim 3; Page 226-228; 1737pp; English.  
 CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic  
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP  
 CC nucleic acids are useful for the production of fine chemicals in  
 CC microorganisms, including organic acids, nonproteinogenic amino acids,  
 CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated  
 CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,  
 CC vitamins, cofactors, polyketides and enzymes  
 XX  
 SQ Sequence 993 BP; 233 A; 256 C; 268 G; 236 T; 0 U; 0 Other;  
 Query Match 41.8%; Score 993; DB 4; Length 993;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-299;  
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1321 CAAAGCAGACCTGTAATGAGATTTCCATGATCACCATCGTGACCTATGGAAGTACTTA 1380  
 Db 1 CAAAGCAGACCTGTAATGAGATTTCCATGATCACCATCGTGACCTATGGAAGTACTTA 60  
 QY 1381 AGTAAATGATTTGTTCTTAACATGTTTAAATATAGTTTCATGACCCCATTTCACTGGA 1440  
 Db 61 AGTAAATGATTTGTTCTTAACATGTTTAAATATAGTTTCATGACCCCATTTCACTGGA 120  
 QY 1441 CACTTTGCTCTCAATCATTCATGATGAGGAGCTTTCGAAGGCGCTCTTACGCTTTCCAT 1500  
 Db 121 CACTTTGCTCTCAATCATTCATGATGAGGAGCTTTCGAAGGCGCTCTTACGCTTTCCAT 180  
 QY 1501 TTCCCTCTCGGCGGTGAGTCAAGCTTTAAAGCTTCTCGACATCACTGAGTGT 1560  
 Db 181 TTCCCTCTCGGCGGTGAGTCAAGCTTTAAAGCTTCTCGACATCACTGAGTGT 240  
 QY 1561 GGTATCGGCGACCCCAAGCGGCAAGCAACCGAGCGGCTGAGTCTTGTGCAAGCAGC 1620  
 Db 241 GGTATCGGCGACCCCAAGCGGCAAGCAACCGAGCGGCTGAGTCTTGTGCAAGCAGC 300  
 QY 1621 GCGGAAATGCTGTTGCTGCAAGCAGAACTAAAGCGCACTATCTGAGCGCTTGTGCA 1680  
 Db 301 GCGGAAATGCTGTTGCTGCAAGCAGAACTAAAGCGCACTATCTGAGCGCTTGTGCA 360  
 QY 1681 AATCCGCTTAACATGCGCATCAACGCGAGTTCGATCCATGATGTTTCTCCGCTGT 1740  
 Db 361 AATCCGCTTAACATGCGCATCAACGCGAGTTCGATCCATGATGTTTCTCCGCTGT 420  
 QY 1741 CAACGAGGTAGCTTCTGGGCTGAGCAACGCTCAGCTGCGCTTGGAGATGAGGCA 1800  
 Db 421 CAACGAGGTAGCTTCTGGGCTGAGCAACGCTCAGCTGCGCTTGGAGATGAGGCA 480  
 QY 1801 CACATATCTCTGCTGCGGCTGAGATGTTTAAAGCGGTAACCCGCTGAGCTAATCC 1860  
 Db 481 CACATATCTCTGCTGCGGCTGAGATGTTTAAAGCGGTAACCCGCTGAGCTAATCC 540  
 QY 1861 CGTGGCGGATGAGTGAATAGTACATGGAACCATCGGCACTTGGCCTATGCAACCC 1920  
 Db 541 CGTGGCGGATGAGTGAATAGTACATGGAACCATCGGCACTTGGCCTATGCAACCC 600  
 QY 1921 CTCAATGCGGATGCTTACATGTTTATGAGGAACTAGATGCGGCTGCGATGCGCTT 1980  
 Db 601 CTCAATGCGGATGCTTACATGTTTATGAGGAACTAGATGCGGCTGCGATGCGCTT 660  
 QY 1981 ACCTTCTGCTCCAAAGATGCTTCAAGACCGGTGACCTGAGCGGCGCTGCGATGCTC 2040  
 Db 661 ACCTTCTGCTCCAAAGATGCTTCAAGACCGGTGACCTGAGCGGCGCTGCGATGCTC 720  
 QY 2041 TGTGGGCGCGAGCGCTATCCATTTGTCGCTGCGGGAAGGTTTGGTGAAGCAATTCG 2100  
 Db 721 TGTGGGCGCGAGCGCTATCCATTTGTCGCTGCGGGAAGGTTTGGTGAAGCAATTCG 780  
 QY 2101 CCGAGCGCTTGGTGGGGAATCTTCTCCGAAACCAAGCTGCTCCCATGCTAAAGCAG 2160  
 Db 781 CCGAGCGCTTGGTGGGGAATCTTCTCCGAAACCAAGCTGCTCCCATGCTAAAGCAG 840  
 QY 2161 AGAAGTATGCTCTCGATGAGATACCATGATGACACCGATGATTTGGCAAGATGCG 2220  
 Db 841 AGAAGTATGCTCTCGATGAGATACCATGATGACACCGATGATTTGGCAAGATGCG 900  
 QY 2221 CCTGGATCTAGATCTCTAGCTAGATCAAGAGCGGCTGCTGATGAGCAATTCAGGG 2280  
 Db 901 CCTGGATCTAGATCTCTAGCTAGATCAAGAGCGGCTGCTGATGAGCAATTCAGGG 960  
 QY 2281 ATTGGCGCTTATGTTACTTCTGAAAAGGTTTCAG 2313  
 Db 961 ATTGGCGCTTATGTTACTTCTGAAAAGGTTTCAG 993

RESULT 5  
 AAS96096  
 ID AAS96096 standard; DNA; 993 BP.  
 XX

AC AAS96096;  
 XX 26-FEB-2002 (first entry)  
 DE C. glutamicum gene #21 encoding metabolic pathway protein.  
 XX Metabolic pathway protein; MP; lysine biosynthesis pathway;  
 KW methionine biosynthesis pathway; large-scale production of fine chemical;  
 XX Corynebacterium diphtheriae; diphtheria; ds.  
 XX Corynebacterium glutamicum.  
 XX WO200166573-A2.  
 XX 13-SEP-2001.  
 XX 22-DEC-2000; 2000WO-IB002035.  
 XX 09-MAR-2000; 2000US-0187970P.  
 PR 23-JUN-2000; 2000US-00606740.  
 XX (BADI ) BASF AG.  
 PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
 PI Kim J, Lee H, Hwang B;  
 XX WPI: 2001-582269/65.  
 DR P-PSDB; AAU71886.  
 XX Nucleic acids encoding metabolic pathway proteins from Corynebacterium  
 PT glutamicum, useful for producing methionine and lysine in Corynebacterium  
 PT and Brevibacterium.  
 XX Disclosure; Page 207-208; 316pp; English.  
 XX The present invention relates to the isolation of novel Corynebacterium  
 CC glutamicum genes encoding metabolic pathways (MP) proteins (AAU71863-  
 CC AAU71922). The metabolic pathway proteins of the invention include  
 CC enzymes involved in the lysine and methionine biosynthetic pathways. The  
 CC polynucleotide sequences of the invention can be used for the large-scale  
 CC production and/or modulation of expression of fine chemicals such as  
 CC lysine and methionine. The sequences of the invention may be used to  
 CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a  
 CC subject to detect diphtheria. AAS96073-AAS96132 represent C. glutamicum  
 CC genes encoding the novel metabolic pathway proteins of the invention  
 XX  
 SQ Sequence 993 BP; 233 A; 256 C; 268 G; 236 T; 0 U; 0 Other;

Query Match 41.8%; Score 993; DB 4; Length 993;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-299;  
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1321 CAAAGCAGACCTGTAATGAGATTTCCATGATCACCATCGTGACCTATGGAAGTACTTA 1380  
 Db 1 CAAAGCAGACCTGTAATGAGATTTCCATGATCACCATCGTGACCTATGGAAGTACTTA 60  
 QY 1381 AGTAAATGATTTGTTCTTAACATGTTTAAATATAGTTTCATGACCCCATTTCACTGGA 1440  
 Db 61 AGTAAATGATTTGTTCTTAACATGTTTAAATATAGTTTCATGACCCCATTTCACTGGA 120  
 QY 1441 CACTTTGCTCTCAATCATTCATGATGAGGAGCTTTCGAAGGCGCTCTTACGCTTTCCAT 1500  
 Db 121 CACTTTGCTCTCAATCATTCATGATGAGGAGCTTTCGAAGGCGCTCTTACGCTTTCCAT 180  
 QY 1501 TTCCCTCTCGGCGGTGAGTCAAGCTTTAAAGCTTCTCGACATCACTGAGTGT 1560  
 Db 181 TTCCCTCTCGGCGGTGAGTCAAGCTTTAAAGCTTCTCGACATCACTGAGTGT 240  
 QY 1561 GGTATCGGCGACCCCAAGCGGCAAGCAACCGAGCGGCTGAGTCTTGTGCAAGCAGC 1620  
 Db 241 GGTATCGGCGACCCCAAGCGGCAAGCAACCGAGCGGCTGAGTCTTGTGCAAGCAGC 300  
 QY 1621 GCGGAAATGCTGTTGCTGCAAGCAGAACTAAAGCGCACTATCTGAGCGCTTGTGCA 1680



QY 2021 GACGGCGCGTGCATGCTCTGCGGCGCAGCGCGTATCATTGTCGCGCGGAA 2080  
 Db 601 GACGGCGCGTGCATGCTCTGCGGCGCAGCGCGTATCATTGTCGCGCGGAA 660  
 QY 2081 GCTTTTGGTGAAGCAATTCGCGAGGCTTGGTGGGAGCTTCTCCGAAACCCAAAGCT 2140  
 Db 661 GCTTTTGGTGAAGCAATTCGCGAGGCTTGGTGGGAGCTTCTCCGAAACCCAAAGCT 720  
 QY 2141 GCTCCCATGCTAAACGAGAGAGTGCATCTCTCGATGAGATACCCATTCACACACCG 2200  
 Db 721 GCTCCCATGCTAAACGAGAGAGTGCATCTCTCGATGAGATACCCATTCACACACCG 780  
 QY 2201 ATGATTGGCAACGATGGCGCTGGAATCTAGATCTCTAGCTAGACTACACAGCGCGTC 2260  
 Db 781 ATGATTGGCAACGATGGCGCTGGAATCTAGATCTCTAGCTAGACTACACAGCGCGTC 840  
 QY 2261 GTTATGACAGCAATCGAGGATTCGGGCT 2290  
 Db 841 GTTATGACAGCAATCGAGGATTCGGGCT 870

RESULT 7

AA71779/C  
 ID AAF71779 standard; DNA; 822 BP.

XX AC AAF71779;

XX DT 30-APR-2001 (first entry)

XX DE Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:53.

XX KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;  
 KW fine chemical production; microorganism; organic acid; nucleoside;  
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;  
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
 KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.

XX OS Corynebacterium glutamicum.

XX PN WC200100843-A2.

XX PD 04-JAN-2001.

XX PF 23-JUN-2000; 2000WO-1B000923.

XX PR 25-JUN-1999; 99US-0141031P.

PR 01-JUL-1999; 99DE-01030476.

PR 02-JUL-1999; 99US-0142101P.

PR 08-JUL-1999; 99DE-01031415.

PR 08-JUL-1999; 99DE-01031418.

PR 08-JUL-1999; 99DE-01031419.

PR 08-JUL-1999; 99DE-01031420.

PR 08-JUL-1999; 99DE-01031424.

PR 08-JUL-1999; 99DE-01031428.

PR 08-JUL-1999; 99DE-01031434.

PR 08-JUL-1999; 99DE-01031435.

PR 08-JUL-1999; 99DE-01031443.

PR 08-JUL-1999; 99DE-01031453.

PR 08-JUL-1999; 99DE-01031457.

PR 08-JUL-1999; 99DE-01031458.

PR 08-JUL-1999; 99DE-01031478.

PR 08-JUL-1999; 99DE-01031510.

PR 08-JUL-1999; 99DE-01031541.

PR 08-JUL-1999; 99DE-01031573.

PR 08-JUL-1999; 99DE-01031592.

PR 08-JUL-1999; 99DE-01031632.

PR 08-JUL-1999; 99DE-01031634.

PR 08-JUL-1999; 99DE-01031636.

PR 09-JUL-1999; 99DE-01032125.

PR 09-JUL-1999; 99DE-01032126.

PR 09-JUL-1999; 99DE-01032130.

PR 09-JUL-1999; 99DE-01032186.

PR 09-JUL-1999; 99DE-01032206.

PR 09-JUL-1999; 99DE-01032227.  
 PR 09-JUL-1999; 99DE-01032228.  
 PR 09-JUL-1999; 99DE-01032229.  
 PR 09-JUL-1999; 99DE-01032230.  
 PR 14-JUL-1999; 99DE-01032922.  
 PR 14-JUL-1999; 99DE-01032926.  
 PR 14-JUL-1999; 99DE-01032928.  
 PR 14-JUL-1999; 99DE-01033004.  
 PR 14-JUL-1999; 99DE-01033005.  
 PR 14-JUL-1999; 99DE-01033006.  
 PR 12-AUG-1999; 99US-0148613P.  
 PR 27-AUG-1999; 99DE-01040764.  
 PR 27-AUG-1999; 99DE-01040765.  
 PR 27-AUG-1999; 99DE-01040766.  
 PR 27-AUG-1999; 99DE-01040832.  
 PR 31-AUG-1999; 99DE-01041378.  
 PR 31-AUG-1999; 99DE-01041379.  
 PR 31-AUG-1999; 99DE-01041380.  
 PR 31-AUG-1999; 99DE-01041394.  
 PR 31-AUG-1999; 99DE-01041396.  
 PR 03-SEP-1999; 99DE-01042076.  
 PR 03-SEP-1999; 99DE-01042077.  
 PR 03-SEP-1999; 99DE-01042079.  
 PR 03-SEP-1999; 99DE-01042086.  
 PR 03-SEP-1999; 99DE-01042087.  
 PR 03-SEP-1999; 99DE-01042088.  
 PR 03-SEP-1999; 99DE-01042095.  
 PR 03-SEP-1999; 99DE-01042124.  
 PR 03-SEP-1999; 99DE-01042129.  
 PR 09-MAR-2000; 2000US-0187970P.  
 XX (BADI ) BASF AG.  
 XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
 PI WPI; 2001-137957/14.  
 XX P-PSDB; AAB79660.  
 DR Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway  
 PT proteins, useful for producing fine chemicals in microorganisms,  
 PT including organic acids, nonproteinogenic amino acids, and purine and  
 PT pyrimidine bases.  
 XX Claim 3; Page 233-234; 1737pp; English.  
 XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic  
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP  
 CC nucleic acids are useful for the production of fine chemicals in  
 CC microorganisms, including organic acids, nonproteinogenic amino acids,  
 CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated  
 CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,  
 CC vitamins, cofactors, polyketides and enzymes  
 XX Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;  
 SQ Query Match 34.6%; Score 822; DB 4; Length 822;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-245;  
 Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 629 GATTCCAAACCCGCGAAACCTTAACCCATCAATCCATCAATCCATCAATCCGCTATC 688  
 Db 822 GATTCCAAACCCGCGAAACCTTAACCCATCAATCCATCAATCCGCTATC 763  
 QY 689 ACACTGCCAGCAGCTTGTATCCAGCCACCTTTGGGGCTTGACAGCGGGGTGAC 748  
 Db 762 ACACTGCCAGCAGCTTGTATCCAGCCACCTTTGGGGCTTGACAGCGGGGTGAC 703  
 QY 749 AATGCTGCTGGCGGAAACCCACAGCGGGAACCAAGATCAGGCTTCCGCGAACCGGCA 808  
 Db 702 AATGCTGCTGGCGGAAACCCACAGCGGGAACCAAGATCAGGCTTCCGCGAACCGGCA 643  
 QY 809 GCGCGAAATCCACCGTCCGGTGTCCCGGTATTTGCGCGCGGCGCGGATAAACACA 868

Db 642 GCGCGAAATCCACCGTCGCTGTCGCGGTAATTCGCGCGCAGCGCGCGGATAAACACA 583  
 QY 869 AACGGTCCAAATACGATTCGGGTTCAACAGGTAGCAGCATTTGCCATCAACATGGGC 928  
 Db 582 AAGCGTCCAAATACGATTCGGGTTCAACAGGTAGCAGCATTTGCCATCAACATGGGC 523  
 QY 929 TTTACCCAAACCCGCTGCTTATCGACGGTCACTCTCAACCGCAGCCGCTTGGCGGTGCA 988  
 Db 522 TTTACCCAAACCCGCTGCTTATCGACGGTCACTCTCAACCGCAGCCGCTTGGCGGTGCA 463  
 QY 989 GTGGCCAGCGCGCAACCGCCAAAGCGTGTCAFGGGAACGGTGTGTTCTTCTTCA 1048  
 Db 462 GTGGCCAGCGCGCAACCGCCAAAGCGTGTCAFGGGAACGGTGTGTTCTTCTTCA 403  
 QY 1049 ATGATCTGTGGCGTCCACCTGTTTGTGTCATGGGCTTTCGTCGTCATGACGCGAAC 1108  
 Db 402 ATGATCTGTGGCGTCCACCTGTTTGTGTCATGGGCTTTCGTCGTCATGACGCGAAC 343  
 QY 1109 CATACAGGTAAAGCGATGCCACCCAGCGCATATATCGAGCAGCATGCGCGCGCATTTG 1168  
 Db 342 CATACAGGTAAAGCGATGCCACCCAGCGCATATATCGAGCAGCATGCGCGCGCATTTG 283  
 QY 1169 GACAAAGATCAACCGCCAGGTGCGCGCATGAAACAAAGACCTCAGAAATTAACAC 1228  
 Db 282 GACAAAGATCAACCGCCAGGTGCGCGCATGAAACAAAGACCTCAGAAATTAACAC 223  
 QY 1229 ACGAGAAACCGCAATGAGTCTTGGCGCTTAATTCCTGTTTAATCACCAGTATTC 1288  
 Db 222 ACGAGAAACCGCAATGAGTCTTGGCGCTTAATTCCTGTTTAATCACCAGTATTC 163  
 QY 1289 TGGGTCGATGAGCAGTAAAGACTGGCCCCCAAGAGACCTGTAAATGAAGATTTCC 1348  
 Db 162 TGGGTCGATGAGCAGTAAAGACTGGCCCCCAAGAGACCTGTAAATGAAGATTTCC 103  
 QY 1349 ATGATCAGCATCGTACCTATGAGAGTACTTAAGTAAATGATGGTCTTCAATGGTT 1408  
 Db 102 ATGATCAGCATCGTACCTATGAGAGTACTTAAGTAAATGATGGTCTTCAATGGTT 43  
 QY 1409 TAATATAGCTTCATGACCCCATTCACCTGGACACTTTGCTC 1450  
 Db 42 TAATATAGCTTCATGACCCCATTCACCTGGACACTTTGCTC 1

RESULT 8  
 AAS96098/c  
 ID AAS96098 standard; DNA; 822 BP.  
 XX AAS96098;  
 AC AAS96098;  
 XX 26-FEB-2002 (first entry)  
 DT C. glutamicum gene #23 encoding metabolic pathway protein.  
 DE  
 DE Metabolic pathway protein; MP; lysine biosynthesis pathway;  
 KM methionine biosynthesis pathway; large-scale production of fine chemical;  
 KW Corynebacterium diphtheriae; diphtheria; ds.  
 XX  
 XX Corynebacterium glutamicum.  
 OS  
 XX W0200166573-A2.  
 PN 13-SEP-2001.  
 PD  
 XX 22-DEC-2000; 2000WO-IB002035.  
 PF  
 XX 09-MAR-2000; 2000US-0187970P.  
 PR 23-JUN-2000; 2000US-00606740.  
 XX  
 XX (BADI ) BASF AG.  
 XX  
 XX Pompejus M, Kroege B, Schroeder H, Zelder O, Haberhauer G;  
 PI Kim J, Lee H, Hwang B;  
 XX

DR WPI: 2001-582269/65.  
 DR P-PSDB; AAU71888.  
 XX Nucleic acids encoding metabolic pathway proteins from Corynebacterium  
 PT glutamicum, useful for producing methionine and lysine in Corynebacterium  
 PT and Brevibacterium.  
 XX Disclosure; Page 214-215; 316pp; English.  
 XX The present invention relates to the isolation of novel Corynebacterium  
 CC glutamicum genes encoding metabolic pathway (MP) proteins (AAU71863-  
 CC AAU71922). The metabolic pathway proteins of the invention include  
 CC enzymes involved in the lysine and methionine biosynthetic pathways. The  
 CC polynucleotide sequences of the invention can be used for the large-scale  
 CC production and/or modulation of expression of fine chemicals such as  
 CC lysine and methionine. The sequences of the invention may be used to  
 CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a  
 CC subject to detect diphtheria. AAS96073-AAS96132 represent C. glutamicum  
 CC genes encoding the novel metabolic pathway proteins of the invention  
 XX  
 SQ Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;  
 Query Match 34.6%; Score 822; DB 4; Length 822;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-245;  
 Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 629 GATTCCAAACCCGCGAAACTAACCCATCAACATCAGTTGATGGCCAAATGGGTCATC 688  
 Db 822 GATTCCAAACCCGCGAAACTAACCCATCAACATCAGTTGATGGCCAAATGGGTCATC 763  
 QY 689 ACAACTGCCACGACGAGTGTGATCCAGCGCCACACCTTTGGGGCTGAGCAGCGGCGTAC 748  
 Db 762 ACAACTGCCACGACGAGTGTGATCCAGCGCCACACCTTTGGGGCTGAGCAGCGGCGTAC 703  
 QY 749 AATGCTGCTGGCGGAAACCCACAGCGGAAACAGATCAGGCTTGGCGGAAACGCGCA 808  
 Db 702 AATGCTGCTGGCGGAAACCCACAGCGGAAACAGATCAGGCTTGGCGGAAACGCGCA 643  
 QY 809 GCGGCGAAATCCACCGTCCGGTGTGCGGCTATTGCGCGCCGACGCCGCCGATTAACACA 868  
 Db 642 GCGGCGAAATCCACCGTCCGGTGTGCGGCTATTGCGCGCCGACGCCGCCGATTAACACA 583  
 QY 869 AACCGTCCAAATACGATTCGGGTTCAACAGGTGAGCAGCATTTGCCATCAACATGGGC 928  
 Db 582 AACCGTCCAAATACGATTCGGGTTCAACAGGTGAGCAGCATTTGCCATCAACATGGGC 523  
 QY 929 TTTACCCAAACCCGCTGCTTATCGACGGTCACTCTCAACCGCAGCCGCTTGGCGGTGCA 988  
 Db 522 TTTACCCAAACCCGCTGCTTATCGACGGTCACTCTCAACCGCAGCCGCTTGGCGGTGCA 463  
 QY 989 GTGGCCAGCGCGCAACCGCCAAAGCGTGTCAFGGGAACGGTGTGTTCTTCTTCA 1048  
 Db 462 GTGGCCAGCGCGCAACCGCCAAAGCGTGTCAFGGGAACGGTGTGTTCTTCTTCA 403  
 QY 1049 ATGATCTGTGGCGTCCACCTGTTTGTGTCATGGGCTTTCGTCGTCATGACGCGAAC 1108  
 Db 402 ATGATCTGTGGCGTCCACCTGTTTGTGTCATGGGCTTTCGTCGTCATGACGCGAAC 343  
 QY 1109 CATACAGGTAAAGCGATGCCACCCAGCGCATATATCGAGCAGCATGCGCGCGCATTTG 1168  
 Db 342 CATACAGGTAAAGCGATGCCACCCAGCGCATATATCGAGCAGCATGCGCGCGCATTTG 283  
 QY 1169 GACAAAGATCAACCGCCAGGTGCGCGCATGAAACAAAGACCTCAGAAATTAACAC 1228  
 Db 282 GACAAAGATCAACCGCCAGGTGCGCGCATGAAACAAAGACCTCAGAAATTAACAC 223  
 QY 1229 ACGAGAAACCGCAATGAGTCTTGGCGCTTAATTCCTGTTTAATCACCAGTATTC 1288  
 Db 222 ACGAGAAACCGCAATGAGTCTTGGCGCTTAATTCCTGTTTAATCACCAGTATTC 163  
 QY 1289 TGGGTCGATGAGCAGTAAAGACTGGCCCCCAAGAGACCTGTAAATGAAGATTTCC 1348  
 Db 162 TGGGTCGATGAGCAGTAAAGACTGGCCCCCAAGAGACCTGTAAATGAAGATTTCC 103

QY 1349 ATGATCACCACCTGACCTATGGAAGTACTTAAGTAAATGATTGGTTCTTAACATGCTT 1408  
DB 102 ATGATCACCACCTGACCTATGGAAGTACTTAAGTAAATGATTGGTTCTTAACATGCTT 43  
QY 1409 TAAATAGCTTCAATGAACCCCAATCACTGGACACTTTGCTC 1450  
DB 42 TAAATAGCTTCAATGAACCCCAATCACTGGACACTTTGCTC 1

RESULT 9

ACC80941/c  
ID ACC80941 standard; DNA; 711 BP.  
XX AC ACC80941;  
DT 27-OCT-2003 (revised)  
DT 11-AUG-2003 (first entry)  
XX  
DE LysE protein encoding sequence.  
KW L-lysine; L-arginine; LysE; ds.  
OS Corynebacterium glutamicum.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..711  
FT /\*tag= a  
FT /product= "lyse protein"  
XX  
FN EPI266966-A2.

XX  
PD 18-DEC-2002.  
XX  
PF 05-JUN-2002; 2002EP-00012539.  
XX  
PR 12-JUN-2001; 2001JP-00177075.  
XX  
PA (AJIN ) AJINOMOTO CO INC.

XX  
PI Gunji Y, Yasueda H;  
XX WPI; 2003-241171/24.  
DR P-PSDB; ABR58213.  
XX

Novel DNA encoding variant of LysE protein from a coryneform bacterium, when introduced into methanol assimilating bacterium, facilitates excretion of L-lysine and/or L-arginine to outside of a cell.

Example 1; Page 17-18; 23pp; English.

The present invention relates to DNA encoding variants of protein with loop region and six hydrophobic helices which facilitates excretion of L-lysine and/or L-arginine to outside of cell of a methanol assimilating bacterium when introduced into the bacterium. The method is used for encoding a protein which facilitates excretion of L-lysine, L-arginine or both of these L-amino acids to outside of a cell of a methanol assimilating bacterium when DNA of the method is introduced into the bacterium. The present sequence represents a lysE protein from Brevibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003 to standardise OS field)

XX  
SQ Sequence 711 BP; 135 A; 173 C; 222 G; 181 T; 0 U; 0 Other;

Query Match 29.9%; Score 711; DB 8; Length 711;  
Best Local Similarity 100.0%; Pred. No. 5,7e-211;  
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 649 CTAACCCATCAACATGAGTTGATGGCCCAATGGCGTCACTCACTGCCACGACGCTT 708  
DB 711 CTAACCCATCAACATGAGTTGATGGCCCAATGGCGTCACTCACTGCCACGACGCTT 652  
QY 709 GATCAGCGGCCACACCTTGGGGCTGGACAGCGGGGTGACATGCTGTGGCGCGAAACC 768

DB 651 GATCAGCGGCCACACCTTGGGGCTGGACAGCGGGCTGACATGCTGTGGCGCGAAACC 592  
QY 769 CACACGCGGGAACACGATACAGGCTTGGCGGAACCGCGCAGCGGGAATAATCCACGCTCC 828  
DB 591 CACACGCGGGAACACGATACAGGCTTGGCGGAACCGCGCAGCGGGAATAATCCACGCTCC 532  
QY 829 GGTGTGCGCGTATTGGCGCGCGGACGCGGATATAACACAAACGCGTCCAAATACGCAAT 888  
DB 531 GGTGTGCGCGTATTGGCGCGCGGACGCGGATATAACACAAACGCGTCCAAATACGCAAT 472  
QY 889 CGGGTTCAACACAGGTGACGACGATTCATCAATGAGGCTTTTACCCAAACCGCGTCTT 948  
DB 471 CGGGTTCAACACAGGTGACGACGATTCATCAATGAGGCTTTTACCCAAACCGCGTCTT 412  
QY 949 ATGACGCTGCTACCTCCACCGGACCGGTTGGCGGTGTGAGTGGCCACCGCGCAACCGCC 1008  
DB 411 ATGACGCTGCTACCTCCACCGGACCGGTTGGCGGTGTGAGTGGCCACCGCGCAACCGCC 352  
QY 1009 CAAAGCGGTGTATCGGGGACCGGTTGGTTCTTCTTCAATGATCTGTGGCGCTTCCAC 1068  
DB 351 CAAAGCGGTGTATCGGGGACCGGTTGGTTCTTCTTCAATGATCTGTGGCGCTTCCAC 292  
QY 1069 CTGTGTTGTGATGCGGTCTTTCGTCGCCATGACGCGCAAAACCAATACAGGTAAGCGATGCC 1128  
DB 291 CTGTGTTGTGATGCGGTCTTTCGTCGCCATGACGCGCAAAACCAATACAGGTAAGCGATGCC 232  
QY 1129 ACCCAGCGCATATATCGGACGATCGCGCGGCGCATTTGGACAAAGAGATCAACGCCAA 1188  
DB 231 ACCCAGCGCATATATCGGACGATCGCGCGGCGCATTTGGACAAAGAGATCAACGCCAA 172  
QY 1189 GGTGCGCGCATGAAACAAAAGACGTGAGAAATTAACACACGAGAAAGAACCGCAATGAG 1248  
DB 171 GGTGCGCGCATGAAACAAAAGACGTGAGAAATTAACACACGAGAAAGAACCGCAATGAG 112  
QY 1249 TCCTTCGCGCTTAATTCCTTGTGTTTAATCAACAGTACATTTCTGGCGTCCGATGACAGTAA 1308  
DB 111 TCCTTCGCGCTTAATTCCTTGTGTTTAATCAACAGTACATTTCTGGCGTCCGATGACAGTAA 52  
QY 1309 AAGACTGGCCCCCAAAAGCAGACCTGTAATGAGATTTCCATGATCACCAT 1359  
DB 51 AAGACTGGCCCCCAAAAGCAGACCTGTAATGAGATTTCCATGATCACCAT 1

RESULT 10

AAH68420/c  
ID AAH68420 standard; DNA; 708 BP.  
XX  
AC AAH68420;  
XX  
DT 26-SEP-2001 (first entry)  
XX  
DE C glutamicum coding sequence fragment SEQ ID NO: 3455.  
XX  
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
XX organic acid synthesis; ds.  
XX  
OS Corynebacterium glutamicum.  
XX  
PN EPI108790-A2.  
XX  
PD 20-JUN-2001.  
XX  
PF 18-DEC-2000; 2000EP-00127688.  
XX  
PR 16-DEC-1999; 99JP-00377484.  
PR 07-APR-2000; 2000JP-00159162.  
PR 03-AUG-2000; 2000JP-00280988.  
XX  
PA (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.  
DR P-PSDB; AAG93201.  
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analyzing  
PT expression profile or pattern of a gene and identifying homologous gene.  
XX Claim 1; SEQ ID NO 3455; 246pp + Sequence Listing; English.  
XX The present invention provides a number of nucleotide and protein  
CC sequences from the Corynebacterium Corynebacterium glutamicum. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of corynebacterium, measuring expression amount and analyzing  
CC the expression profile or expression pattern of a gene derived from  
CC Corynebacterium, and identifying a homologue of a gene derived from  
CC Corynebacterium. Corynebacterium bacteria are useful for producing amino  
CC acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a nucleic acid described  
CC in the exemplification of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the European Patent Office  
XX  
SQ Sequence 708 BP; 134 A; 173 C; 221 G; 180 T; 0 U; 0 Other;  
Query Match 29.8%; Score 708; DB 5; Length 708;  
Best Local Similarity 100.0%; Pred. No. 4.9e-210;  
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 652 ACCGATCAATCATGTTGATGCGGCAATGCGGTATCATCACTGCCAGGACGATGAT 711  
DB 708 ACCGATCAATCATGTTGATGCGGCAATGCGGTATCATCACTGCCAGGACGATGAT 649  
QY 712 CCAGCCCAACCTTTGGGCTGCACAGCGGGCTGACATGCTGCGCGGAAACCCAC 771  
DB 648 CCAGCGGCAACCTTTGGGCTGCACAGCGGGCTGACATGCTGCGCGGAAACCCAC 589  
QY 772 CAGCGGAAACCATGATGAGGCTTCCGCGAAGCGCGGCGGAAATTCACCGTCGGT 831  
DB 588 CAGCGGAAACCATGATGAGGCTTCCGCGAAGCGCGGCGGAAATTCACCGTCGGT 529  
QY 832 GTCCGCTATTGCGGCGGACGCGCGGATTAACACAGAGCGTCCAAATACGATTCGG 891  
DB 528 GTCCGCTATTGCGGCGGACGCGCGGATTAACACAGAGCGTCCAAATACGATTCGG 469  
QY 892 GTTCAACAGGTCAGCAGATTCGCAATCAATGCGGCTTACCAAAACCGCTGCTATC 951  
DB 468 GTTCAACAGGTCAGCAGATTCGCAATCAATGCGGCTTACCAAAACCGCTGCTATC 409  
QY 952 GACGCTCACCTCCACCGCACCGGTTGCGGTGTCAGTGGCGCACCGCGCAACCGCCAA 1011  
DB 408 GACGCTCACCTCCACCGCACCGGTTGCGGTGTCAGTGGCGCACCGCGCAACCGCCAA 349  
QY 1012 AGCGGTGTCATCGGCGAGGTTGGTCTGTTCTTCAATGATCTGCGGCTTCCAGCTT 1071  
DB 348 AGCGGTGTCATCGGCGAGGTTGGTCTGTTCTTCAATGATCTGCGGCTTCCAGCTT 289  
QY 1072 GTTTGTTCATGCGGCTTTTCGCTGCCATGACGCGCAACCAATACAGGTAAAGTGCAC 1131  
DB 288 GTTTGTTCATGCGGCTTTTCGCTGCCATGACGCGCAACCAATACAGGTAAAGTGCAC 229  
QY 1132 CCAGCGGATAATATCAGCAGCATCGGCGGCGATTTGGACAAAGATCAACGCGCCAGGT 1191  
DB 228 CCAGCGGATAATATCAGCAGCATCGGCGGCGATTTGGACAAAGATCAACGCGCCAGGT 169  
QY 1192 GCCGGCGATGAACAAAGAGCGTCAAAATTAACACAGAGAAACCGCAATGATGCC 1251  
DB 168 GCCGGCGATGAACAAAGAGCGTCAAAATTAACACAGAGAAACCGCAATGATGCC 109  
QY 1252 TTCGCGCTTAATTCCTTTTAAATCACCAGTACATTTCTGGGTCGATGAGCAGTAAAG 1311  
DB 108 TTCGCGCTTAATTCCTTTTAAATCACCAGTACATTTCTGGGTCGATGAGCAGTAAAG 49

QY 1312 ACTGGCCCCCAAAAGCAGACCTGTATGAGATTTCATGATCACCAT 1359  
DB 48 ACTGGCCCCCAAAAGCAGACCTGTATGAGATTTCATGATCACCAT 1  
RESULT 11  
AAH45375/c  
ID AAH45375 standard; DNA; 1568 BP.  
XX AAH45375;  
AC  
XX 11-SEP-2001 (first entry)  
DT  
XX C. thermosynogenes lysin biosynthetic enzyme lyse DNA.  
DE  
XX Heat-resistant; lysin biosynthesis; enzyme; coryneform;  
KW aspartate-semialdehyde dehydrogenase; lyse; ds.  
KW  
OS Corynebacterium thermosynogenes.  
XX  
PN JP2001120270-A.  
XX  
PD 08-MAY-2001.  
XX  
PF 01-NOV-1999; 99JP-00311148.  
XX  
PR 01-NOV-1999; 99JP-00311148.  
XX  
PA (AJIN ) AJINOMOTO KK.  
XX  
XX WPI; 2001-364760/38.  
DR P-PSDB; AAG64047.  
XX  
PT A heat-resistant lysin biosynthetic system enzyme gene of a high  
PT temperature-resistant coryneform microbe.  
XX  
PS Example 5; Page 22-24; 27pp; Japanese.  
XX  
CC The invention relates to a gene from a high temperature-resistant  
CC coryneform microbe that encodes a heat-resistant lysin biosynthetic  
CC enzyme. The enzyme has aspartate-semialdehyde dehydrogenase activity and  
CC can be used for growing amino acid-producing microbes. The present  
CC sequence encodes an enzyme of the invention  
XX  
SQ Sequence 1568 BP; 258 A; 525 C; 490 G; 295 T; 0 U; 0 Other;  
Query Match 29.4%; Score 698.6; DB 4; Length 1568;  
Best Local Similarity 68.4%; Pred. No. 7e-207;  
Matches 1017; Conservative 0; Mismatches 454; Indels 16; Gaps 3;  
QY 650 TAACCCATCAACATCAGTTTGTATGGCCAAATGCGGTATCACTGCGACGACGCTTG 709  
DB 1472 TAACCCATCAGGATCAGCTTCAAGGCAATCCGCTGAGCACCACGCGCCATATGTTG 1413  
QY 710 ATCCAGCGCCACACCTTGGGCTGACAGCGGGCTGCAATGCTGCGCGGAAACCC 769  
DB 1412 ATCCAGCGCCACACCTTGGGCTGACAGCGGGCTGACAGCGGGCTGCGCGCGCTAACCG 1353  
QY 770 ACCAGCGGGAACACAGATCAGGCTTGGCGGAAACCGCGGCGGCAAAATCCACCGTCCG 829  
DB 1352 ACCAGTGGGAACACAGGCTTGGCGGGAAGGACCCCGCAGGAAGATCCACCGACCG 1293  
QY 830 GTGTGCGCGGTATGCGCGCGACGCGCGCGATTAACACAAACGCGTCCAAATAGCGATTC 889  
DB 1292 GTCTCCCGGTACTGCGGCTCCGACACCGGATGAAGACGAAGGATCCAGGTAGGATTC 1233  
QY 890 GGGTTCAACAGGTGACAGGATTTGCCATCACTGAGGCTTTACCCCAACCGCTGCTTA 949  
DB 1232 GAGTTGAGCAGGTGACGACAAATGGCGCATGAGCATGAGGCTTGACCCAGCTTCGCGGTG 1173  
QY 950 TCGACGCTCACCTCCACCGCGACCCCGGTTGCGGCTGTGTCAGTGGCCACCGCGGAAACCGCC 1009  
DB 1172 CCTGATGTGATCGGAGCGCGGGTCTGTTGTTCTGCTCAGCGCCCGCGGAGGCGGAC 1113









AC AAF68077;  
 DT 11-APR-2001 (first entry)  
 DE Corynebacterium glutamicum MCT protein encoding DNA SEQ ID NO:669.  
 XX  
 KW Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;  
 KW membrane construction and membrane transport protein; petroleum spill;  
 KW hydrocarbon degradation; gram positive aerobic bacterium; marker;  
 KW identification; microorganism; fine chemical production; transformation;  
 KW genome mapping; genetic engineering; ds.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN WO200100805-A2.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 23-JUN-2000; 2000WO-1800926.  
 XX  
 PR 25-JUN-1999; 99US-0141031P.  
 PR 08-JUL-1999; 99DE-01031454.  
 PR 08-JUL-1999; 99DE-01031478.  
 PR 08-JUL-1999; 99DE-01031563.  
 PR 09-JUL-1999; 99DE-01032122.  
 PR 09-JUL-1999; 99DE-01032124.  
 PR 09-JUL-1999; 99DE-01032125.  
 PR 09-JUL-1999; 99DE-01032128.  
 PR 09-JUL-1999; 99DE-01032180.  
 PR 09-JUL-1999; 99DE-01032182.  
 PR 09-JUL-1999; 99DE-01032190.  
 PR 09-JUL-1999; 99DE-01032191.  
 PR 09-JUL-1999; 99DE-01032209.  
 PR 09-JUL-1999; 99DE-01032212.  
 PR 09-JUL-1999; 99DE-01032227.  
 PR 09-JUL-1999; 99DE-01032228.  
 PR 09-JUL-1999; 99DE-01032229.  
 PR 09-JUL-1999; 99DE-01032300.  
 PR 14-JUL-1999; 99DE-01032327.  
 PR 14-JUL-1999; 99DE-01033005.  
 PR 14-JUL-1999; 99DE-01033006.  
 PR 27-AUG-1999; 99DE-01040764.  
 PR 27-AUG-1999; 99DE-01040765.  
 PR 27-AUG-1999; 99DE-01040766.  
 PR 27-AUG-1999; 99DE-01040830.  
 PR 27-AUG-1999; 99DE-01040831.  
 PR 27-AUG-1999; 99DE-01040832.  
 PR 27-AUG-1999; 99DE-01040833.  
 PR 31-AUG-1999; 99DE-01041378.  
 PR 31-AUG-1999; 99DE-01041379.  
 PR 31-AUG-1999; 99DE-01041395.  
 PR 03-SEP-1999; 99DE-01042077.  
 PR 03-SEP-1999; 99DE-01042078.  
 PR 03-SEP-1999; 99DE-01042079.  
 PR 03-SEP-1999; 99DE-01042088.  
 XX  
 PA (BADI ) BASF AG.  
 XX  
 PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
 XX  
 DR WPI; 2001-071486/08.  
 DR P-PSDB; AAB76844.  
 XX  
 PT Corynebacterium glutamicum nucleic acids encoding membrane construction  
 PT and membrane transport proteins or their portions, useful for typing or  
 PT identifying C. glutamicum or related bacteria, and as markers for  
 PT transformation.  
 XX  
 PS Claim 3; Page 1107-1108; 1119pp; English.  
 XX  
 CC AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane  
 CC construction and membrane transport (MCT) proteins given in AAB76510 to  
 CC AAB76847. The MCT nucleic acids and proteins are useful in the

CC identification of microorganisms which can be used to produce fine  
 CC chemicals, for modulating fine chemical production in C. glutamicum or  
 CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or  
 CC identification of C. glutamicum or related bacteria, as reference points  
 CC for mapping C. glutamicum genome, and as markers for transformation.  
 CC AAF68082 and AAF68082 represent sequencing primers which are used in an  
 CC example from the present invention  
 XX  
 SQ Sequence 993 BP; 222 A; 247 C; 277 G; 247 T; 0 U; 0 Other;  
 Query Match 20.4%; Score 485.2; DB 4; Length 993;  
 Best Local Similarity 99.0%; Pred. No. 2.7e-140;  
 Matches 499; Conservative 0; Mismatches 3; Indels 2; Gaps 1;  
 QY 1 AGATACTCTTTTGAAGAAACCATGTACGATTGCGTGACATTGTTCGCTCGAAAGGC 60  
 DB 490 AGATACTCTTTTGAAGAAACCATGTACGATTGCGTGACATTGTTCGCTCGAAAGGC 549  
 QY 61 TCTTTACGTGGGTATTTCTTCTTACGCTCCAGAGCTCACAGCGAGGCGGCTGAGTTCA 120  
 DB 550 TCTTTACGTGGGTATTTCTTCTTACGCTCCAGAGCTCACAGCGAGGCGGCTGAGTTCA 609  
 QY 121 GCGGAGGAGGGCTGCGCGGCTTCTGATTTCATCAGCAAGCTATTCATCATTAATCGTTG 180  
 DB 610 GCGGAGGAGGGCTGCGCGGCTTCTGATTTCATCAGCAAGCTATTCATCATTAATCGTTG 669  
 QY 181 GGTGGAGGAACCGGGCGATGACGCTGAGAACTTTGTCAGTCAGCTGCCAACAAATTCGATGG 240  
 DB 670 GGTGGAGGAACCGGGCGATGACGCTGAGAACTTTGTCAGTCAGCTGCCAACAAATTCGATGG 729  
 QY 241 TGGCGTCAATTGCTTTCTTCTTCCACCTTGGCGAGGCGCTGCTCACGCAAAATATCTCGATGG 300  
 DB 730 TGGCGTCAATTGCTTTCTTCTTCCACCTTGGCGAGGCGCTGCTCACGCAAAATATCTCGATGG 789  
 QY 301 AATTCCAGAGGGTTCCCGCGCCAGCCAGGCTAGTCCCTGCTGAGGGGCTATTTGAAGCT 360  
 DB 790 AATTCCAGAGGGTTCCCGCGCCAGCCAGGCTAGTCCCTGCTGAGGGGCTATTTGAAGCT 849  
 QY 361 GAACAATATTGATATGGTCCGCAAGCTCAATGACATCGCCAGCAAGCGGGGAGTCACT 420  
 DB 850 GAACAATATTGATATGGTCCGCAAGCTCAATGACATCGCCAGCAAGCGGGGAGTCACT 909  
 QY 421 TGGCGAGATGCGGCTTGCATGCGGCTGCGGAGGCAAGGAGAGTACGCGGC--GGATACC 478  
 DB 910 TGGCGAGATGCGGCTTGCATGCGGCTGCGGAGGCAAGGAGAGTACGCGGC--GGATACC 969  
 QY 479 GTGACCACTGCATTGATTGGTCT 502  
 DB 970 GTGACCACTGCATTGATTGGTCT 993

RESULT 15  
 ACA29651  
 ID ACA29651 standard; DNA, 1095 BP.  
 XX  
 AC ACA29651;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Prokaryotic essential gene #11308.  
 XX  
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 XX  
 OS Corynebacterium diphtheriae.  
 XX  
 PN WO200271183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 XX (ELIT-) ELITRA PHARM INC.  
 XX  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 DR P-PSDB; ABU25781.  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 14; SEQ ID NO 17521; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway;  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1095 BP; 264 A; 247 C; 311 G; 273 T; 0 U; 0 Other;

Query Match 12.9%; Score 305.8; DB 7; Length 1095;  
 Best Local Similarity 67.1%; Pred. No. 3.5e-84;  
 Matches 433; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

QY 1 AGATATCTCTTTGGAGAAACCATGTACGCATTGCGTGACATTGTTGGTCTGGAAGGC 60  
 DB 450 AGATACGCCCTTTGGAGAAACCATGTATGCTTTGCGAGACATGTTGGCATCCGGCAAGGC 509  
 QY 61 TCTTTACGTGGGTATTTCTTCTTCCAGAGCTCACAGCGGAGCGGCTGAGTTTCA 120  
 DB 510 TCGTATGTGGGAATTTCTCTTATGACCTGAACCGCAGAGATCTCGGAATTTCT 569  
 QY 121 GCGGAGGAGGGCTCCCGCTTCTGATTTCATCAGCAAGCTATTCATCATTAATCGTTG 180  
 DB 570 AGCTGTGTAGGGCTCCCGCTGCTGATTTCATCAACCGAGCTATTTATCTTGAACCGCTG 629  
 QY 181 GGTGAGGAGACCGGCGATACGGTGAGACTTTTTCAGTCAGTCAGTCACACATGTTCT 240  
 DB 630 GGTAGAGAAACCAAGGTGAAGTGGCGAGAACTGTTAGAGTCTGCGGCGGAATAACGGACT 689

QY 241 TGGCTCATTTGCTTTCTCACCATTTCGCGAGGCGCTGCTCAGGACAAATATCTCGATGG 300  
 DB 690 GGGTGTATTGCAATTTTCGCCACTTTGACACAGGGGCTTTTAACTGACAGGTATCTCGAGG 749  
 QY 301 AATTCCAGAGGGTTCCCGGCCAGCCAGCGTAAGTCCCTGCTCTGAGGGCATGTTGAACGT 360  
 DB 750 CGTGCCAGTGGATTTCAGGCGCAGCAGCAGGGAAGTCTCTTGGGCAAGAGATGCTCAACGC 809  
 QY 361 GAACATATTGATGATGTCGCGAAGCTCAATGACATGCGCCAGGAAACCGGCAAGTCACT 420  
 DB 810 TAAAAATCTTGACATGCTGCGTGCCTCACTGAATGATTAGCGTTACGACGTGCCCAACGTT 869  
 QY 421 TGGCAGATGGCGCTTCATATGGTGTCTGCGCAGCAGGAGAGTACGGCGCGGATACCGT 480  
 DB 870 GGCACAGCTGGCTATTGCTGGTGTCTGCGGACAGGCGGATTTATGGGCCCAACGCT 929  
 QY 481 GACCAGTGCATTTGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 DB 930 AACAGTGGCTTGTATCGGAGCATCGTTCGTAGCCAGTTGGATCAAAATCTGGGAGCTCT 989  
 QY 541 CAACAACTTGGAGTTTCTGACGCCAGTTCGAGCGGATGCGATGAGATTTCACACGACGC 600  
 DB 990 TAATATCTTGAATTCAGCGCTTGAGGAAACCGCAATATATTGACACTGTTGCCAAGGACGC 1049  
 QY 601 CGGCATCAACATTTGGGCGGAAGGCCACCGCATTTCCAAAACCCGCGA 645  
 DB 1050 TGGATTTAATATTTGGGCGGCTGCTACTGCTCGAAGGTTTCAGGA 1094

Search completed: March 15, 2004, 14:59:19  
 Job time : 971.046 secs



1493 CTTTCCATTTCCCTCGCGGTGAGTCAGCGCGTTAAAGCTCTCGAGCATCAGCGGT 1552  
 5313 CTACATGTCACCCCTCGCTGTCAGTCAGCGCATCAAGTCGTTGGAGCAGCAGTCGCG 5372  
 1553 CGAGTGTGTTGATCGCGCACCACCGGCAAAAGCAACCGAAGCGGGTGAAGTCTTGTG 1612  
 5373 CAGGTGCTGTTGTCAGGAAAGCCATGTCGGCGGACGACCGCAGGTATCCGCTGTTG 5432  
 1613 CAGCAGCGCGGAAATGTTGTCGAGCAGAACTAAAGCGCACTATCTGG--AC 1670  
 5433 CGGTTGCGCGCAACAGCGGTTGCTCGAGTCCGAGCGCTCGCTGNAATGGGTGGCAAC 5492  
 1671 GCCTTGTCTGAA-----ATCCGTTAAACATCGCCATCAACGAGATTCCTATCCACATCG 1726  
 5493 GCGTCGCTGAAACGCGCGGATCACCATTGCGGTAAACGCGATTCATCGGCGCATGG 5552  
 1727 TTTCTCTCGCTGTTACGAGGTAGCTTCTTGGGGTGAGCAACGCTCACGCTCGCGTTG 1786  
 5553 TTTTCGCGCGTGTTCGACG-----GTCTCGCGCAGCTCTCTGCTCGAGTTGCGATC 5603  
 1787 GAAGATGAAGCGCACACATATCTTGTGCGCGGTGAGATGTTTGAAGCGGTAAAC 1846  
 5604 GAGACACAGACCATTCGCGCGCTGCTACGCGAGGTTGCGGATGGGCGCGGTGAC 5663  
 1847 CGTGAAGCTAAATCCGCGCGGATGTGAAGTAGTAGAACTTGGAAACATCGCCACTTG 1906  
 5664 ACCGAGCGGNAACCGGTGCGCGCTCGCGGTGACCGCTGGGTGMAATGCGTACCTA 5723  
 1907 GCMATGCAACCCCTCATTTGCGGATGCTACATGTTGATGGGAACTAGATGGGT 1966  
 5724 CCAATGCGCAGCAGGCAATTCGTCAGCGCCATCTATCCGAGCGGTTCACTGCGCGCG 5783  
 1967 GCGATGCGCGTCTTACGCTGCGTCCCAAGATGTCTTCAAGACGCTGACCTGCGACGG 2026  
 5784 GCGCTAAAGCTCCGCTGCTGCGTGGATGGAATCGTGAATGCGTGGCGTGAATGTTG 5843  
 2027 CGC---GTCCGATGTCCTGTTGGGCGCAGGCGGTATCCATTTGCTCCGTCGCGGAAGT 2083  
 5844 CGTAAGGCTTTGTCGCGCCATCACAGACCGACGACCTTTGTCGCGACACACAGAGGCG 5903  
 2084 TTTGGTAGGCAATTCGCGGAGCTTGTGGGAGCTTCTTCCGAAACCCAGCTGCT 2143  
 5904 TTAACCGCGCGACGCGCGCGCTGGGATGGGCAATGTTCCCGGAGAAGCTGGCAGCA 5963  
 2144 CCAATGTAAGACGAGGAGATGATCTCTCGATGAGATACCAATGACACACCGATG 2203  
 5964 TCTCCGCTGCGGATGATGCTGCTGCGGCTGCGGACATACACTCGAGCTCTCTC 6023  
 2204 TATTTGCAACGATGCGCGCTGGAATCTAGATCTAGCTAGACTACAGACCGCTGTT 2263  
 6024 TATTTGCAATGCTGGAATGGAAGTGGACAGTCGATCATCGCGGAATACCGACACGTTGAG 6083  
 2264 GATCAGCAATCAGGAGTTGCGGCTTAGTTAC 2297  
 6084 GCGCGGCAAGCGCTGCTGATACCGGCGCGCAAC 6117

RESULT 2  
 US-09-894-844-12  
 ; Sequence 12, Application US/09894844  
 ; Patent No. 6686166  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Behr, Marcel  
 ; APPLICANT: Small, Peter  
 ; APPLICANT: Schoolnik, Gary  
 ; APPLICANT: Wilson, Michael A.  
 ; TITLE OF INVENTION: Molecular Differences Between Species of  
 ; TITLE OF INVENTION: the M. Tuberculosis Complex  
 ; FILE REFERENCE: STAN102CON  
 ; CURRENT APPLICATION NUMBER: US/09/894,844  
 ; CURRENT FILING DATE: 2001-06-27  
 ; PRIOR APPLICATION NUMBER: 09/318,191  
 ; PRIOR FILING DATE: 1999-05-25

; PRIOR APPLICATION NUMBER: 60/097,936  
 ; PRIOR FILING DATE: 1998-08-25  
 ; NUMBER OF SEQ ID NOS: 137  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 12  
 ; LENGTH: 909  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacteria tuberculosis  
 US-09-894-844-12

Query Match 5.1%; Score 121.2; DB 4; Length 909;  
 Best Local Similarity 49.5%; Pred. No. 5.9e-26;  
 Matches 433; Conservative 0; Mismatches 423; Indels 18; Gaps 4;

Qy 1433 CAACCTGACACACTTCTCTCAATCATTTGATGAAGCGAGCTTCGAAGCGCTCTCTTAGCC 1492  
 Db 28 CAGCTGCGCGCATTTGGCTGCGGTGTCGAACTGGGCGAGCTTCGATGCGGCGCGGAGCGC 87  
 Qy 1493 CTTTCCATTTCCCTCGCGGTGAGTCAGCGGTAAAGCTCTCGAGCATCAGTGGGT 1552  
 Db 88 CTACATGTCACCCGCTGCGCTGTCAGTCAGCGCATCAAGTCGTTGGAGCAGCAGTCCGC 147  
 Qy 1553 CGAGTGTGATTCGCGCACCCAAACCGGCCAAAGCAACCGGCTGAAGTCTTGTG 1612  
 Db 148 CAGTGTGCTGCTGTCAGGAAAGCCATGTCGCGCGACGACCGCAGGTATCCCGCTGTTG 207  
 Qy 1613 CAAGCAGCGCGGAAATGCTTGTCTGCAAGCAGAACTAAAGCGCACTATCTGG--AC 1670  
 Db 208 CGTGTGCGCGCAACAGCGTGTCTGAGTCCGAGGCGCTCGCTGAAATGGGTGGCAAC 267  
 Qy 1671 GCCTTGTCTGAA-----ATCCGTTAAACATGCGCATCAAGCAGATTCGCTATCCACATGG 1726  
 Db 268 GCGTGTGCTGAAACGCGCATTCGCGGTAAACCGCATTCCTATGCGCGCACATGG 327  
 Qy 1727 TTTCTCTCGCTGTTCAAGCGGTAGCTTCTTGGGCTGAGCAACGCTCAACGCTGCGCTTG 1786  
 Db 328 TTTTCGCGCGTGTGTGACG-----GTCTCGCGGAGCTTCTCTGCTGCACTGCGGATC 378  
 Qy 1787 GAAGATGAAGCGCACACATATATCTTCTGCTGCGGCTGGAGATGTTTGAAGCGCTTAAC 1846  
 Db 379 GAGGACCGAGCCATTCGCGCGCTGCTACGCGAGGCTGTGCGGATGCGCGCTGACC 438  
 Qy 1847 CGTGAAGCTAATCCGCTGCGGAGTGAAGTAGTAGAACTTGGAAACCATGCGCATGG 1906  
 Db 439 ACCGAGCGGAAACCGGCTGCGGCTGCGGCTGCAACCGCTGGGTGAAATGCGCTACTA 498  
 Qy 1907 GCAATTCACACCCCTCAATTCGCGGATGCTACATGTTGATGGGAACTAGATTTGGCT 1966  
 Db 499 CAGTGTGCGCAGCGGCAATTCGCTCCAGCGCATCTATCCGAGCGGTTCATGCGCGCGG 558  
 Qy 1967 GCGATGCGCTCTTACGCTTCGCTCCCAAGATGCTTCAAGACCGTCACTGAGCGGG 2026  
 Db 559 GCGGCTAAAGCTCCGCTCACTGCGGTGGAATCTGACGATGGGCTGCGAGCATGTTGGTG 618  
 Qy 2027 CGC---GTCCGATGCTGTTGGGCGCAGGCGGTATCCATTTGCTCCGCTGCGGAGGT 2083  
 Db 619 CGTAAGGCTTTGTCGCGCCATCACAGACCGACGCACTTTGTCCCGACACACAGAGGCG 678  
 Qy 2084 TTTGCTGAGGCAATTCGCGAGGCTTGTGTTGGGACCTTCTTCCCGAAACCCAGCTGCT 2143  
 Db 679 TTAACCGCGCAGCGGCGCGGCTGGGATGGGCACTGTTCCCGGAGGTGGCAGCA 738  
 Qy 2144 CCAATGCTAAAGCAGGAGAGTGTCTCTCTCGATGAGATACCCATTGACACACCGATG 2203  
 Db 739 TCTCCGCTTCCGATGATGCTGTCGACGCGTCTCGGACATACACCTCGACGCTCCCTC 798  
 Qy 2204 TATTTGCAACGATGCGGCTGGAATCTAGATCTTAGCTAGACTACAGACCGCTGTT 2263  
 Db 799 TATTTGCAATGCTGGAATGGAAGTGGACAGTCGCGGGAATACCGACACGCTGAGG 858  
 Qy 2264 GATGCGCAATCGAGGATTCGCGCTTAGTTAC 2297  
 Db 859 GCGCGGCAAGCGCTGCTGATACCGGCGCGCAAC 892

### RESULT 3

```

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2 4403765
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          5.1%; Score 121.2; DB 3; Length 4403765;
Best Local Similarity 49.5%; Pred. No. 9.8e-24;
Matches 433; Conservative 0; Mismatches 423; Indels 18; Gaps 4;

QY      1433  CAACGTGGACATTTGGCTCTCAATCATTTGATGAAGCAGCTTCGAAGCGCTCTCTTAGCC 1492
Db       |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
QY      1493  CTTTCCATTTCCCCCTCGCGCGTAGTCAGCGCGTTAAAGTCTCTCGAGCATCAAGTGGGT 1552
Db       |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
QY      1553  CGAGTGTGGTATCGCGACCCACCGCGCCAGCAACCGAAGCGGGTGAAGTCTTTGGT 1612
Db       |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
QY      1613  CAAGCAGCGCGGAAATGTGTGCTCAAGCAGAGAACTAAAGCGCAACTATCTGG--AC 1670
Db       |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
QY      1671  GCCTTGCTGAA----ATCCGTTAAACATCGCCATCAACGAGANTTCGTATCCACATGG 1726
Db       |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
QY      1727  TTTCTCCCGGTGTTCAACAGGTAGCTTCTTGGGGTGGAGCAACGCTCAACGCTCGCGCTG 1786
Db       |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
QY      1787  GAAGATGAAGCGCAACATATTCTTGTCTCGCGCGTGGAGATGTTTGGAGAGCGGTAAAC 1846
Db       |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
QY      1847  CGTGAAGCTATATCCCGTGGCGGATGTGAAGTAGTAGAATTCGAAACATTCGCGCACTGG 1906
Db       |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
QY      1907  GCCATTGCAACCCCTCATTTGGGGATGCCCTACATGGTTGATGGGAAACTAGATTGGGCT 1966
Db       |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
QY      1967  GCGATGCCCGCTTTACGCTTCGGTCCCAAGATGTGCTTCAAGCGGTGACTTGACGGGG 2026
Db       |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
QY      2027  CGC---GTGATGGTCTGTGGGGCGGACGGCGGTATCCATTTGTCCTCGTGGCGGAGGT 2083

```

## RESULT 4

```

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-200007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match	5.1%	Score 121.2	DB 3	Length 4411529
Best local Similarity	49.5%	Pred. No. 9.8e-24		
Matches 433	Conservative 0	Mismatches 423	Indels 18	Gaps 4
QY 1433	CAACTGGACATTTTGCTCTCAATCATTTGATGAAGCGAGCTTTGGAAGCGCGCTCTTTAGCC	1492		
Db 2229873	CAGCTGGCGCGATTTGGCTGCGGTGTGCAACTGGCGAGCTTCGATGCGCGCGCGGAGCGCG	2229814		
QY 1493	CTTTTCATTTTCCCCCTCGCGCGGTGAGTCAGCGCGGTAAAGCGTCTCGAGCATCATCGTGGGT	1552		
Db 2229813	CTACATGTCAACCGGTGCGCTGTCTAGTCAGCGCATCAAGTCGTTTGAGCAGCAGGTCCGC	2229754		
QY 1553	CGAGTGTGTGTATCGCGCACCCACCGGCCAAGCAACCGAAGCGGGTGAAGTCTCTGTG	1612		
Db 2229753	CAGTGTCTGTGTGTTCAGGGAAAGCCATGTCTGGGCGACACCGCAGGTATCCGCGTGTG	2229694		
QY 1613	CAAGCAGCGCGGAAATGTGTCTCTGCAAGCAGAACTAAAGCCCACTATCTGG--AC	1670		
Db 2229693	CGGTTGGCGCGCAACAGCGTGTCTGAGTCCGAGCGCTCGCTGAAATGGGTGGCAAC	2229634		
QY 1671	GCCTTGCTGAA-----ATCCCGTTAAACATCGCGCATCAACGCAAGATTCCGTATCCATCG	1726		
Db 2229633	GCCTCGCTGAACGACGCGGATCAACATTTGGGTAAACCGCGATTCATGGGCACTGG	2229574		
QY 1727	TTTTCTCCCGGTGTTCAACGAGGTAGCTTCTTGGGGTGGAGCAACGCTCAGCGTGGCTTG	1786		
Db 2229573	TTTTCGGCGGTGTTGACG-----GTCTCGGCGAGCTCTCTGCTCGAGCTTCGATC	2229523		
QY 1787	GAAAGATGAAGCGCACACATATCTCTGCTGGCGGTGGAGATGTTTTAGGAGCGGTAAACC	1846		

Db 2229522 GAGGACGAGGACCAATTCGGCGGGCTGCTACGGAGGGTGTGGCGATGGGGCGGTGACC 2229463  
 QY 1847 CGTGAAGCTAATCCGGTGGCGGAGTGAAGTGTAGAACTTGGAAACCATGGCCCACTTG 1906  
 Db 2229462 ACCGAGCGGAACCCGGTGGCGGGCTGCCGGGTGACCCCGTGGGTGAATGGCTACCTA 2229403  
 QY 1907 GCATTTGCAACCCCTCATTTGGCGGATGCTACATGGTGTGATGGGAACTAGATTGGCT 1966  
 Db 2229402 CCAGTGGCCAGCAGCCATTCGTCAGCGCCATCTATCCGAGCGGTTCACTGGCCCGCGG 2229343  
 QY 1967 GCGATGCCGCTTACGCTTCGGTCCCAAGATGTCTTCAAGACCGGTGACCTGGACGGG 2026  
 Db 2229342 GCGGCTAAGCTCCGCTCACTGGCGTGGAACTGTCAGATGGGCTCGAGCAATGTTGGTG 2229283  
 QY 2027 GCG---GTGATGTCTGTGGGGGCGAGGCGGTATTCATTTGTCCTGGGGAAGT 2083  
 Db 2229282 CGTAAGCGCTTTGTCGCGCCATCACACAGCCGACGCTTTGTCTCCGACACAGAGGCG 2229223  
 QY 2084 TTTGGTGAGGCAATTCGCGAGGCTTGGTGGGACTTCTCCCGAAACCAAGCTGCT 2143  
 Db 2229222 TTACGCGCGAGCGCGCGGCTGGGATGGGCAATGTTCCCGAGAGCTGGAGCA 2229163  
 QY 2144 CCCATGCTAAAGCAGGAGAGTGTATCTCTCGATGAGATACCCATTTGACACACCGATG 2203  
 Db 2229162 TCTCCGCTTCCGATGATGCTGCTGAGGGTCTGGACATACACTCGAGCTCCCTTC 2229103  
 QY 2204 TATTTGGAAGATGGCGCTGGAACTAGATCTCTAGTAGACTCACAGAGCGCGTCT 2263  
 Db 2229102 TATTTGCAATGCTGGAACATGAGACAGTCCGATCATCGCGGAATACCGACAGCGTGAGG 2229043  
 QY 2264 GATGACAGCAATCGAGGATTCGCGCTTGTAGTTAC 2297  
 Db 2229042 GCGCGGCAAGCGGTCTGTACCGGGGCGGACGAC 2229009

RESULT 5  
 US-08-920-812-20/c  
 ; Sequence 20, Application US/08920812  
 ; Patent No. 5763188  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ohno, Tsuneya  
 ; APPLICANT: Matsuhisa, Akio  
 ; APPLICANT: Uehara, Hirotugu  
 ; APPLICANT: Eda, Soji  
 ; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
 ; NUMBER OF SEQUENCES: 25  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States of America  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/920,812  
 ; FILING DATE: 29-AUG-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION INFORMATION:  
 ; APPLICATION NUMBER: US 08/362,577  
 ; FILING DATE: 27-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Rin-Laures, Li-Hsien  
 ; REGISTRATION NUMBER: 33,547  
 ; REFERENCE/DOCKET NUMBER: 19036/32420  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312/474-6300  
 ; TELEFAX: 312/474-0448  
 ; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 5541 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: Genomic DNA  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Escherichia coli  
 ; STRAIN: Clinical Isolate EC-625  
 ; US-08-920-812-20  
 Query Match 4.2%; Score 100.8; DB 1; Length 5541;  
 Best Local Similarity 50.5%; Pred. No. 2.2e-19;  
 Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;  
 QY 1 AGATPACTCTTTTGGGAAGAACCAATGTACGCATGTGGTGA CATTGTTGGTCTTGGAAAGGC 60  
 Db 3416 AATATCGCCGATGGAAGAAACCGCTCTCGCTGGCTCATGCGGTACAAAGCGGTAAAGC 3357  
 QY 61 TCTTTAGCTGGGTATTTCTTCTACGGTCCAGAGCTCAGAGCGGAGCGGCTGAGTTTCAT 120  
 Db 3356 GCTTTATGTGCGGATCTCTCTTACTCGCCAGAGCGGACGCAAAAATGTTGAGTTGCT 3297  
 QY 121 GCGCGAGGAGGCTCGCGCTTCTGATTCTATCAGCAAGCTATTCCATCATTAATCGTTG 180  
 Db 3296 GCGGAGTGGAAATTCGCTGTGTTAATTCATCACTTCGTACAAATTACTGAACCGCTG 3237  
 QY 181 GGTGAGGAACCGGGGATGAGCGGTGAGAACTTTGTTGAGTCTGAGTCCCAACATGCTCT 240  
 Db 3236 GGTGGATAAAGCGG-----CCTGTGATACCTTCGAAAAATAACGCGCT 3192  
 QY 241 TGGCGTCAATGCTTTCTCACCCTTGGCAGGCGCTCTCAGGACAAATATCTCGATGG 300  
 Db 3191 GGGCTGTATTGCTTTACTCTCTGGCTAGGGATTTCTGCCGGAATATCTCAACGG 3132  
 QY 301 AATTCCAGAGGTTCCCG-----CCGAGCGAGGTAAGTCCCTGCTCTGAGGG 348  
 Db 3131 CATTCGCAAGATTCAACGATGATCGTGAAGGGAATAAAGTTCTGTTGCTGACCGGAA 3072  
 QY 349 CATGTTGAAGTGAACATATTGATGTTGCTCGCAGCTCAATGATCATGCCAGGACG 408  
 Db 3071 AATGCTTACCGAAGCAACCTCAACAGCCTTACCTTATTGAATGAATGGCACAGCG 3012  
 QY 409 CGGCGAGTCACTTGGCGAGATGGCGCTTGCATGGGTGCTGCGGAGCAAGGAGGATACGG 468  
 Db 3011 TGGACATCAATGCGCGCAATGGCGTTAAGCTGGTTGCTGAAGA-----2967  
 QY 469 CGCGGATACCGTGACCAAGTGTGATGTTGTTGCTTCTGCTGAGTTCAGCAGCTGACACAG 528  
 Db 2966 ---TGATCGGCTGACGTGCGTATTGATTGGTGC-CAGCGCGCGGACAACTTTGAGGAGAA 2911  
 QY 529 CTTTGATTCACTCAAACTTTGAGTCTTCTGACCGCGAGTTGGAGCGCATCGATGAGAT 588  
 Db 2910 CTTGAGCGCTGATTAATCTGACATTTAGCACCGAGGAGCTGGCCAGATTGATCAGCA 2851  
 QY 589 TTCCCGACGACCGCGCATCAACATTTGGGCGAAGCCACCGATTTCCAAACCCCGGAAAA 648  
 Db 2850 TATCGCGATGGCGAGCTGAATCTGTGCGAGGCGTCTTCCGATAAATGACCTGTTAATA 2791  
 QY 649 CTAAACCATCAACATCAGTTTGTGCGCAATGCGGTCTATCACAAGTCCACGACGAGTT 708  
 Db 2790 CGGCGCGGAGAACCGCGCGCCGATTAAGATCAGTGTATTGACGAGTCAAGTCCGTCAGAT 2731  
 QY 709 GATCCAGCGCCACACCTTTGGGCGTGGACAGCGCGGCTGACAAATG 752  
 Db 2730 AGCCCATCAACAATGACAGAGACCAAAAGTCAAGTGTGATGATG 2687

RESULT 6  
 US-08-920-827-20/c  
 ; Sequence 20, Application US/08920827  
 ; Patent No. 5770375

```

;
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; LENGTH: 5541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: Clinical Isolate EC-625
; US-08-920-827-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;
Best Local Similarity 50.5%; Pred. No. 2.2e-19;
Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

QY 1 AGATACCTCTTTGGAGAAACCATGTACGCAITGGGTGACATTTGGCGTCTGGAAAGGC 60
DB 3416 AAATACGCCGATGGAAGAACCGCTCTGGCTGCTCATGCGGTACAAAGCGGTAAAGC 3357
QY 61 TCTTTAGTGGTATTTCTTCTACGTCCAGAGTCCAGCGGAGCGGCGGTGATTCAT 120
DB 3356 GCTTTATGTCCGATCTCTCTTACTCTCGCAGAGCGGACGCAAAAATGGTTGCT 3297
QY 121 GCGCAGAGGCGTCCCGCTTCTGATTCATCAGCAAGCTATTCCATCATTAATCTTG 180
DB 3296 GCGCGAGTGAATAATTCGCTGTTAATTCATCACTTCGTCATTAATTCGACCGGT 3237
QY 181 GGTGAGGAAACCGGCGATGACGGTGAGAACTTTGTCAGTCAGTCGCAAAATGGTCT 240
DB 3236 GGTGATATAAGCGG-----CCTGCTGATACCTCGCAAAATACGGCGT 3192
QY 241 TGGGCTCATTTGCTTTCTACACATTCGCGAGGCGCTGCTCAGGCAAAATATCTGATGG 300
DB 3191 GGGCTGATTTGGCTTTTACTCTCTCTGCTCAGGGAATTTGCTGACCGGAAATATCTCAACGG 3132
QY 301 AATTCAGAGGGTTTCCG-----CGCCAGCCAGGGAAGTCCCTCTCTGAGGG 348
DB 3131 CATTCGCAAGATTCAGGATGCACTGTAAGGGAATAAAGTTGCTGTCAGCCGAA 3072

```

```

;
; RESULT 7
; US-08-921-177-20/c
; Sequence 20, Application US/08921177
; Patent No. 5798211
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921,177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```







ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-5370

Query Match 3.4%; Score 81.4; DB 4; Length 1095;  
Best Local Similarity 51.5%; Pred. No. 5.2e-14;  
Matches 252; Conservative 0; Mismatches 216; Indels 21; Gaps 2;

QY 1 AGATATCTCTTTGGAAGAACCATGATGACATTTGGTGTGCTGCGTGAAGGC 60  
DB ACAGACCGCTGATGGAACGATGGTGGCTGATCATCTGGTTCTCAGGTTAAGC 587  
QY 61 TCTTTACGTGGGTATTTCTTACCGTCCAGAGCTCACAGCGAGGCGGCTGAGTTCA 120  
DB GCTATATGTGCGGATCTCCAACTATCTCTGCGCCAGCGCGGAGGCGGTGAAGATCCT 647  
QY 121 GCGGAGGAGGCGTCCGCTTCTGATTCATCAGCAAGCTATTCATCATTAATGTTG 180  
DB TAACGACCTCGGCACCCCTGATCATTCACCGCTCTGCTACTCGATGTTTCGAACGCG 707  
QY 181 GGTGAGGAACCGGCGATGACGCTGAGAACTTGTTCAGTCAAGCTGCGCAACATGCT 240  
DB CTGGAAGAGGA-----TTGCTGGATTTTCTGCAGCGGAGGAT 749  
QY 241 TGGGCTCATGTTTTCACACTTCCGACGCGCTGCTCACGGAACAATATCTCGATGG 300  
DB TGAAGCATAGCGTTTTCACCGCTGCGCGGCGCACTTACCGACCGCTATCTCAACGG 809  
QY 301 AATTCNAGAGGTTCCCG---CGCAGCCAGGTTAAGTCCCTGTCTGAGGSCATGTTGA 357  
DB CATTCGCGCGGACTCGCGCGCGCCAGCAGCAGTCTGTTCTGCGAGCGGAGCACTGAC 869  
QY 358 CGTGAACAATATGATATGTTGCGCAAGCTCAATGACATCGCCAGGAACCGCGGAGTC 417  
DB TCCAGCCGATTAGAAAATTCGTGAGTTTACAGGCGGAGGCGCTGGCGAGAA 929  
QY 418 ACTTGCAGATGGCGTTGATGCGGTGCTGCGGAGCAAGAGAGTACGCGCGGATAC 477  
DB GCTGTCGAGATGGCGTGGCTGGCTGGTGTGCGGAGAGAAATGATCCTCGTGTGT 989  
QY 478 CGTGACCAG 486  
DB CGGCGCCAG 998

RESULT 11

US-09-489-039A-4735  
; Sequence 4735, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709,2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 4735  
; LENGTH: 915  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-4735

Query Match 3.4%; Score 80.2; DB 4; Length 915;  
Best Local Similarity 44.8%; Pred. No. 1.1e-13;  
Matches 396; Conservative 0; Mismatches 478; Indels 9; Gaps 2;  
QY 1423 GAACCCCATCACTGGACATTTGCTCTCAATCATTTGATGAAGGCGAGCTTCGAAGCGC 1482  
DB GGACTACAGACACTACAGCGCTGGATGGGTGATTAGGGAACGAGGATTTGAGCGCG 92  
QY 1483 CTCCTTAGCCCTTTCATTTCCCTCGCGGCTGAGTCAGCGGTTAAAGCTCTCGAGCA 1542

DB 93 CGCGCAAGAGCTATGATTTATCCAGTCCGCTCTCAAGCTATTAGCAGCTGGAATA 152  
QY 1543 TCACGTGGTTCGAGTGTGTTGATGCGCACCCAGCAAGCAAGCAGCGGCTGA 1602  
DB TATGTTGCGCCAGCGCTGCTGTTGCTGATACGTCGCGCGCGCCGCCAGCAGGACA 212  
QY 1603 AGTCCTTTGCAAGCAGCGCGGAAAATGTTGCTGCAAGCAGAAACTTAAAGCGCACT 1662  
DB GAAACTGCTGGCTCTGTTGCGCAGGTTGAATCTGAGGAGAGTGGCTGGCGGATGA 272  
QY 1663 ATCTGAGCGCTTGTGTAATTCGCTTAAACATGCGCATCAAGCAGATTCGCTATCCAC 1722  
DB ACAAAACCGGCTCTAGCGGCTTGTGCTGCTGCGGCTGAACGCGCAGACTCTGGCGAC 332  
QY 1723 ATGCTTCTCCCGGTTCAACGAGGTAGCTTCTTGGGSGTGAGCAAGCTCAAGCTGG 1782  
DB CTGCTGCTGCGCGGCTGGCGCAACGTTCTGTAGATCCCTTATTCGTCTCAACCTGCA 392  
QY 1783 CTTGGAAGATGAAGCGCAACATTTATCTTGTCTGCGCGTGGAGATGTTTATGAGCGGT 1842  
DB GGTGAGATGAACCCGACCCAGGAGCGCTGCGGCTGGCGAGTAGTGGCGCAGT 452  
QY 1843 AACCGGTGAAGCTAATCCCGTGGGAGTGTGAAGTAGTAGAATTGGAACCATCGGCCA 1902  
DB GAGTATCCAGCGCGAGCGGCTGCGCAAGCTGCTGTGTGATCAGCTGGGCGGCTCGATTA 512  
QY 1903 CTTGCGCATTGCAACCCCTCTATTCGCGGATGCTTACATGCTTGTGAGGAACTAGATTG 1962  
DB CTTGTTGCTGCMATAAAGAGTTTGGCAACGCTATTTCCGAAATGGCGTGAACGCTTC 572  
QY 1963 GCTGCGAT-----GCCGCTTTACGCTTCCGTTCCCAAGATGTCGTTCAAGACCGTGA 2016  
DB GCGCTGCTTTAAGCGCGCGCTGCTGCGCTTCGACCATCTGACGATATGTCATCAGCGCTT 632  
QY 2017 CTTGCAAGCGCGGCTGATGCTGCTGTTGGGCGGAGGCGGATCATTGTCGCTGGCG 2076  
DB CTTGCAAGCAAACTTCGACCTTGGCGCGGCGAGCGTCCCTTGGCCACATCGTCAACTGCTC 692  
QY 2077 GGAAGGTTTGTGAGGCAATTCGCGAGGCTTGGTTGGGAGCTTCTTCCGAAACCCA 2136  
DB GGAAGCTTTGTCGAGCTGCGCGCCAGGCGCACCATCTGCTGTATGATCCCGCATCTGCA 752  
QY 2137 AGCTGCTCCCATCTAAAGCAGGAGAGTGTCTCTCTGATGAGATACCATTTGACAC 2196  
DB GATCGAAAAAGAGCTCAACAGCGCGAACTTATCGATCTCACCGCGGCTGTTCCAGCG 812  
QY 2197 ACCGATG---TATTGGCAACGATGCGGCTGGAATCTAGATCTCTAGCTAGACTCAGA 2253  
DB CGCATGCTCTACTGGCACCGCTTTCGCGCGGAAAGCGCATGATGCGCGGCTGACCGA 872  
QY 2254 CGCGCTGTTGATGACGCAATCGAGGATTTGGCGCTTTAGTTA 2296  
DB TCGCTCAATTGACTACGACACAGGTTGCTGCTCAGGATTA 915

RESULT 12

US-09-543-681A-791  
; Sequence 791, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRAE  
; FILE REFERENCE: 2709,1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 791  
; LENGTH: 936  
; TYPE: DNA

ORGANISM: Proteus mirabilis  
US-09-543-681A-791

Query Match 3.1%; Score 73.4; DB 4; Length 936;  
Best Local Similarity 44.9%; Pred. No. 1.2e-11;  
Matches 324; Conservative 0; Mismatches 391; Indels 6; Gaps 1;

QY 1455 TCATTGATGAAGGCGAGCTTCGAAGCGGCTCTTAGCCCTTTCCATTTCCTCCGCGG 1514  
DB 80 TTTATCGAGAGCGTGTGTTTGAACGAGCGGCAAAAATATGATTACGCAATCTGCG 139  
QY 1515 TGAGTCAGCGGTTAAAGCTCTCGAGCATCACTGGGTCGAGTGTGGTATCGCCACCC 1574  
DB 140 TTTCAAGAAGTAATAACAACCTTGAGATCTATTTCGACAACCTTTATTAGTACGTACGG 199  
QY 1575 AACCGCCCAAGCAAGCGGTCGAGTCTCTGTCGAGCGCGGCAAAATGTTGT 1634  
DB 200 TTCGCGCCCAACCCACAGAGAGCGGGCAAAAGCTCTGCAATTAATCATAGGTGAA 259  
QY 1635 TGCTGCAAGCAGAACTAAAGCGCACTATCTGAGCGCTTCTGTAATCCGTTAAACA 1694  
DB 260 TGTAGAGAGCAGTGTAGGTGATGAATAATAGTGTCTGCGCGCTCTTTACTCTTT 319  
QY 1695 TCGCATCAACGAGATCGCTATCCATGTTTCTCCGCTGTTCAACGAGTAGTT 1754  
DB 320 TAGCTGTCAACGCGGATGTTAGCTACTTGGTTATTGCTGCTTTCACCGGTATTA 379  
QY 1755 CTTGGGTTGAGCAAGCTCAGCTGCGCTTGAAGATGAAGCGCACACATTAATCTTGC 1814  
DB 380 CACAACTTCTATCCGCTTAATAATCAAGTAGAATGAATGAATCTGTAAGAGCGAT 439  
QY 1815 TGCGCGTGGAGATGTTTATGAGCGGTAAACCGTGAAGCTAAATCCGCTGGCGGATGTG 1874  
DB 440 TAAGACGTGTGAGTGTGTTGCTATCAGTATTCAGCCCAAGCATTCGCTAGCTGTC 499  
QY 1875 AAGTATGAGACTTGAACCATGCGGCATCTGGCCATTTGCAACCCCTCATTTGCGGATG 1934  
DB 500 TAGTCGACCAATAGCGCATTAGACTATCTTTTGTGCAATCCCTGATTTTGCCCAAC 559  
QY 1935 CTTACATGTTTGTAGTGG3-----AACTAGATTGGCTGGCATCCGCTTACCTTCG 1988  
DB 560 GTATTATTGCAATGGGGTCACTAAATCACTCTGTTAAAGCCCGCTGCGCATTTG 619  
QY 1989 GTCCAAAGATGTCTTCAAGACCGTGACCTGCAAGCGGCGGTGATGTCCTGCGGCG 2048  
DB 620 ACCATCTTGATGATGATGATCAGGCTTCTTACAGCAAACTTCGGTTTATCTCCGGA 679  
QY 2049 GCAGCGCGTATCATGTGTCGCGGGAAGGTTTGGTGAAGCAATTCGCGAGGCT 2108  
DB 680 GTGTGCGCTGCCATATTGTTAAATCTTCCGAGCGCTTTGTTCAAGTAAACAAGCT 739  
QY 2109 TTGGTTGGGACTTCTTCCGAAACCAAGCTGCTCCCATCTAAAGCAGGAGTGA 2168  
DB 740 CACATGTTGATGATCCCTCATCTACAAATTCGATGAGCTAAAGTGGTGAATGA 799  
QY 2169 T 2169  
DB 800 T 800

RESULT 13

US-09-252-991A-7644  
Sequence 7644, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 7844  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7644

Query Match 2.7%; Score 65.2; DB 4; Length 900;  
Best Local Similarity 46.2%; Pred. No. 3.2e-09;  
Matches 415; Conservative 0; Mismatches 458; Indels 25; Gaps 5;

QY 1418 TTCAATGAACCCCATTAACATGAGACACTTGTCTCTCAATCATTTGATGAGGCGGCTTCGAA 1477  
DB 3 TTGTTTCGACTACAGTTGCTCGCGCGCTGCGCGAGTATCGACCGGATCAAGCTGCTC 62  
QY 1478 GGGCGCTCTTAGCCCTTTTCCATTTCCCTCGCGGCTGAGTCAGCGGTTAAAGCTCTC 1537  
DB 63 CGCGCGCACAGGCGCTGGGCTGTGCAATCGGCGGTATCGACCGGATCAAGCTGCTC 122  
QY 1538 GAGCATCAGTGGGTGAGTGTGATTCGCGCAACCCAAACCGGCAACGAGCAACGAGCG 1597  
DB 123 GAGCGCGGCTCGCGCAGCGGCTCTGTTGCGGAGACCGCGCATCCCAACGATCTC 182  
QY 1598 GGTGAAGTCTTGTGCAAGCAGCGGAAATGTTGTTGTCAGCAGCAAACTAAAG-- 1655  
DB 183 GGGCGCGCTGTTTCAACCATGTCAGAGTGGGCTGCTGAGGGGAGCTTGCAGCGC 242  
QY 1656 -----CGCAACTATCTGAGCGCTTCTGAAATCCCTTAACCATCGC-CATCAACGC 1707  
DB 243 TGGGTGCGCAACTGATGAAGCGGCTGCCCGGAGCGCTGCGCATCGTGTCAACGC 302  
QY 1708 AGATTGCTATTCACATGTTTCTCCGTTTCAAGGAGTGTCTTGGGGTGGAGC 1767  
DB 303 CGATAGCTTGGCACTTGGTGGCGCGCGGTTGCGGATTTCTGCGCGAGCGGCGGT 362  
QY 1768 AAGCTTCACGCTGCGCTTGAAGATGAAGCGCACACATTAATCTTCTGCTGGCGGTGAGCA 1827  
DB 363 GCTGCTGAGCACTTGTGTGAGGACAGAGGTGGGCTCAAGCGCATGCGCGGCGCA 422  
QY 1828 TGTGTTAGAGCGGTAAACCGTGAAGCTTAATCCGTTGGCGGAGATGTGAAGTAGTAAGT 1887  
DB 423 AGTGGCGGCTGCTCTGCGGTAGCGCGCGGCTGGCGCGGCTGCTGCTGCTGCT 482  
QY 1888 TGCAACCATGCGGCACCTTGGCCATTTGCAACCCCTCATTT-----GGGGATGCTACAT 1941  
DB 483 CGAGCCATGCGTTATCGCGGCTTGGCCAGCCCGGATTTTCAATGCGCGGACCTTTTCCCG 542  
QY 1942 GGTTCATGAGGAACTAGATTGGGCTGCGATGCGCGCTTTACCGCTTCCGATCCCAAGATGT 2001  
DB 543 CGGGTTCAGGCGGCGCTTGGCGCGGCTGCGCGGATCGTGTTCGCGCGGACGCT 602  
QY 2002 GCTTCAAGACGCTGACCTGAGACGGGCGGCTGATGTTCTGFGGGGCGCAGGCGGCTATC 2061  
DB 603 GCTTCAACACCGCTTTCTCAAGGACCTCGCG-----TCGAGGCGGCTTTTTCATCCACCA 656  
QY 2062 CATTTGTCGCGCGGAGAGGTTTGGTGGGCAATTCGCGGAGGCTTGGTGGGAGCT 2121  
DB 657 CTTTGGCGGCTTCTGGAAGGTTTCTGCGGCTTCCCGCGGCTTCTCGGCTGGGCGCT 716  
QY 2122 TCTTCCCGAAACCCCAAGCTGCTCCCATGCTAAAGCAGGAGAGTGAATC---CTCCGCA 2178  
DB 717 GGTTCGCGGAGCGGCTGAGTGAAGGCGAGCTGGCGCGGCTGAACTGTGCGAGCTGCTGCC 776  
QY 2179 TGAGATACCCATTGACACCGATGTTATGGCAACGATGGCGCTTGAATCTTAGATCTCT 2238  
DB 777 CGGCCAGGTCTATGACGATGATCCGCTTGTATGCGCACTACTGCGCAACGCGGCGAACTGT 836  
QY 2239 AGCTAGACTCAGACCGCTGTTGATGAGCAATTCGAGGATTTGGGCTTAGTTA 2296  
DB 837 CGGTGCTTACCAGGACCTCTCTGCGCGCGCGGAGCGCTTGGTGGGCTGTCA 894

RESULT 14

US-09-724-623-24  
 ; Sequence 24, Application US/09724623  
 ; Patent No. 6476209  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glenn, Matthew  
 ; APPLICANT: Lubbers, Mark W  
 ; APPLICANT: Dekker, James  
 ; TITLE OF INVENTION: Polynucleotides, materials incorporating  
 ; TITLE OF INVENTION: them, and methods for using them.  
 ; FILE REFERENCE: 1048U1  
 ; CURRENT APPLICATION NUMBER: US/09/724,623  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 124  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 24  
 ; LENGTH: 1002  
 ; TYPE: DNA  
 ; ORGANISM: Lactobacillus rhamnosus  
 US-09-724-623-24

Query Match 2.7%; Score 64; DB 4; Length 1002;  
 Best Local Similarity 47.9%; Pred. No. 7.8e-09;  
 Matches 279; Conservative 0; Mismatches 270; Indels 33; Gaps 2;  
 QY 9 CTTTGAAGAAACCATGTACGCAATTGCGTGACATTTGCGTCTGGAAGGCTCTTTACG 68  
 DB 446 CTTTGAAGAGACGGTTAATGCGTGATCAAAACGGTGACGGTAAGCGGTTGATA 505  
 QY 69 TGGGTATTTCTTCTTACGGTCCAGAGCTCACAGCGGAGGCGGTGAGTTTCATGCGGAG 128  
 DB 506 TTGGTATTTTCAACTATGATACGAAGCAGACCAAGAGCAATTCGATGTTTAAAGATC 565  
 QY 129 AGGCGTCCCGCTTCTGATTCATCAGCCAGCTTATTCATCATTAATCGTTGGGTGAGG 188  
 DB 566 TGCACACGCGCTTTGTACTGAATCAATACAGTTTAAATCGCACGCTGAAA 625  
 QY 189 AACCGGCGATGACGGTGAGAACTTCTGCGAGTCAAGTGCACCAATCGTCTTGGCGTCA 248  
 DB 626 GGTCCGG-----CTTGATCGATGATTAAGAGCTGATGTCGCGGTTGA 570  
 QY 249 TTGCTTTCTCACACTTGGCGAGGCGCTGCTCAGCGACAAATATCTGATGGAATTCAG 308  
 DB 671 TTGCATACGACCGGTTATCAGAGGCTTGTATCAGATCGCTACCTAAAGGGAATTCGG 730  
 QY 309 AGGTTTCCCGCCAGCCAGCGGTAAGTCCCTGTCTGAGGCGATGTTGAACGTGAACATA 368  
 DB 731 ATACTTTCAAAATCCATCCAAACAGGCCACTTTTGTAGGGCAAGAGGCTGTGG 790  
 QY 369 TTGATATGCTCCGCAAGCTCAATGACATCGCCAGGAACCGGGCAGTCACTTGGCGAGA 428  
 DB 791 TTAAGCAACTAAATCGGTTAATGAAATTTGCGATGATCTGACCAACCTGAGTCAAA 850  
 QY 429 TGGCGCTTGATGGGTGCTGCGGAGCAAGAGAGTACGGCGCGGATACCGTGACCAAGTG 488  
 DB 851 TGGCGCTTGGCGGTGTTTACGGGATCCGG-----TTGTCACAAAGTG 892  
 QY 489 CATTTGATGGTCTTCTGTCAGTTGAGCAGCTGGCAACAGCCTTGTGATTCACCAACT 548  
 DB 893 TGATCATTTGGAGAGACCTTCAGTTGAACACCTTCAGATTAACCTTAAAGCAACGACATC 952  
 QY 549 TGGAGTTTTCTGACCGCGAGTTGGAGCGGATCGATGAGATTT 590  
 DB 953 TGACCTTTACTGCTGAAGAGATTCAACAAATTTGATGATTTT 994

RESULT 15

US-09-252-991A-7908/c  
 ; Sequence 7908, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 7908  
 ; LENGTH: 699  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-7908

Query Match 2.6%; Score 61.8; DB 4; Length 699;  
 Best Local Similarity 54.6%; Pred. No. 2.9e-08;  
 Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;  
 QY 650 TAACCCATCAACATCAGTTTGTATGCGCAATGCGGTCTCATCAAACTGCCACGACGAGCTTG 709  
 DB 698 TATCCCGGAAACAGCAGTGTGGCGGCGCATGCCAGCATCATGGCCGCCACCATCAGGTGG 639  
 QY 710 ATCCAGCGCCACACCTTGGGGCTGGACAGCGGGCGTGACAACTCTGCTGCGCCGAAACCC 769  
 DB 638 AGCAGGCGCCAGGTGGCGGCGCGCCAGCCAGAGGGGCGACCATGCGCGCCGAGGGCG 579  
 QY 770 ACCAGCGGGAACAGATCAGGCTTGGCGGAGACGCGGCGGCGGCAAAATCCACCGTCCG 829  
 DB 578 AGGCGGAAGAACCAATCAGGAGGCGCTGGCGCGCGCGGAGGCGCATAGGCGCGCGCGCG 519  
 QY 830 GTGTGCGCGGTATTGGCGCGCGAGCGCGCGGATAAACACAAACCGCTCCAAATACGCAATC 889  
 DB 518 G-----CCTGCTGGCGCGCGAGCGACCGATCAGAGTAGCGGTATCGAGATAGACGTGG 465  
 QY 890 GGGTTCAACAGGTGACGACGATTTGCCATCA 920  
 DB 464 GGGTTGAGCAGGGGTGACCGCCAGTGGCGCCA 434

Search completed: March 16, 2004, 03:01:32  
 Job time : 197.837 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2004, 00:51:03 ; Search time 851.071 Seconds  
(without alignments)  
10269.549 Million cell updates/sec

Title: US-09-105-117K-3

Perfect score: 2374  
Sequence: 1 agatactctttggaagaa.....gttaacattcagcaaatgg 2374

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	2374	100.0	2374	14	US-10-196-232-24
2	2374	100.0	3309400	9	US-09-738-626-1
3	993	41.8	993	10	US-09-746-660A-47
4	870	36.6	870	9	US-09-738-626-3456
C 5	822	34.6	822	10	US-09-746-660A-51
C 6	711	29.9	711	14	US-10-166-142-7
C 7	708	28.8	708	9	US-09-738-626-3455
C 8	696.8	28.4	712	14	US-10-166-142-9
9	627	26.4	627	9	US-09-738-626-3454
10	485.2	20.4	993	12	US-10-627-476-669
11	305.8	12.9	1095	12	US-10-282-122A-17521
12	302.8	12.8	879	12	US-10-282-122A-17523
13	164.4	6.9	1026	15	US-10-369-493-35374
14	164.4	6.9	1026	15	US-10-369-493-38826
15	164.4	6.9	1028	15	US-10-369-493-38124

16	132.6	5.6	1041	12	US-10-282-122A-13437	Sequence 13437, A
17	123.8	5.2	897	12	US-10-282-122A-14749	Sequence 14749, A
18	122.8	5.2	909	12	US-10-282-122A-26429	Sequence 26429, A
19	121.2	5.1	909	9	US-09-894-844-12	Sequence 12, Appl
20	121.2	5.1	909	15	US-10-388-902-12	Sequence 12, Appl
21	121.2	5.1	912	12	US-10-282-122A-28444	Sequence 28444, A
22	120.8	5.1	954	14	US-10-156-761-1110	Sequence 1110, A
23	120.8	5.1	9025608	14	US-10-156-761-1	Sequence 1, Appl
24	116.4	4.9	1041	12	US-10-282-122A-11887	Sequence 11887, A
25	115.2	4.9	993	14	US-10-156-761-1151	Sequence 1151, Ap
C 26	115.2	4.9	9025608	14	US-10-156-761-1	Sequence 1, Appl
27	114.8	4.8	1107	12	US-10-282-122A-14390	Sequence 14390, A
28	108.6	4.6	1026	15	US-10-369-493-37651	Sequence 37651, A
29	108.6	4.6	1038	12	US-10-282-122A-19965	Sequence 19965, A
30	107	4.5	1041	12	US-10-282-122A-7020	Sequence 7020, Ap
31	105.2	4.4	999	12	US-10-282-122A-38764	Sequence 38764, A
32	105.2	4.4	999	12	US-10-282-122A-39389	Sequence 39389, A
33	97.4	4.1	1035	12	US-10-282-122A-33379	Sequence 33379, A
34	93.2	3.9	988	9	US-09-974-300-77	Sequence 77, Appl
35	86.4	3.6	745	15	US-10-260-238-1450	Sequence 1450, Ap
36	82.6	3.5	894	9	US-09-815-242-6195	Sequence 6195, Ap
37	82.6	3.5	894	12	US-10-282-122A-20492	Sequence 20492, A
38	81.4	3.4	999	12	US-10-282-122A-12754	Sequence 12754, A
39	79.8	3.4	909	12	US-10-282-122A-41572	Sequence 41572, A
40	78.2	3.3	990	12	US-10-282-122A-23370	Sequence 23370, A
C 41	78	3.3	1077	9	US-09-738-626-1395	Sequence 1395, Ap
42	77.4	3.3	990	12	US-10-282-122A-41645	Sequence 41645, A
43	76.2	3.2	894	9	US-09-815-242-9644	Sequence 9644, Ap
44	76.2	3.2	894	12	US-10-282-122A-40108	Sequence 40108, A
45	75.2	3.2	894	12	US-10-282-122A-38717	Sequence 38717, A

## ALIGNMENTS

RESULT 1  
US-10-196-232-24/c  
; Sequence 24, Application US/10196232  
; Publication No. US2003013899A1  
; GENERAL INFORMATION:  
; APPLICANT: YAMAGUCHI, MIKIO  
; APPLICANT: ITO, HISAO  
; APPLICANT: GUNII, YOSHIYA  
; APPLICANT: YASUEDA, HISASHI  
; TITLE OF INVENTION: METHOD FOR PRODUCING L-ARGININE  
; FILE REFERENCE: 225391USO  
; CURRENT APPLICATION NUMBER: US/10/196,232  
; CURRENT FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: JP 2001-224586  
; PRIOR FILING DATE: 2001-07-25  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 2374  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1025)..(1723)  
; OTHER INFORMATION:  
US-10-196-232-24

Query Match	100.0%	Score 2374;	DB 14;	Length 2374;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2374;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AGTACTCTCTTGGAGAACCATGTACGATTCGTCGACATTTGTCGCTCGAAGGC	60	
Db	2374	AGATACTCTCTTGGAGAACCATGTACGATTCGTCGACATTTGTCGCTCGAAGGC	2315	
QY	61	TCTTTACGTGGGTATTTCTTCTACGTCACAGAGCTCACAGCGAGCGCGCTGAGTTTCAT	120	
Db	2314	TCTTTACGTGGGTATTTCTTCTACGTCACAGAGCTCACAGCGAGCGCGCTGAGTTTCAT	2255	



Qy	421	TGCGCAGATGCGCGCTTGCATGCGGTCTGCGCGACGAGGAGTACGCGCGGATACCGT	480
Db	1328015	TGCGCAGATGCGCGCTTGCATGCGGTCTGCGCGACGAGGAGTACGCGCGGATACCGT	1328074
Qy	481	GACCACTGCATTGATTTGGTGTCTTCAGTTGACGAGCTGGACAACAGCGCTTGATTCAT	540
Db	1328075	GACCACTGCATTGATTTGGTGTCTTCAGTTGACGAGCTGGACAACAGCGCTTGATTCAT	1328134

## RESULT, T. 2.

QY	541	CACCAACTTGGAGTTCTTCGACCGCGAGTTGAGCGCATCGATGAGATTCTCCACGACGC	600
Db	1328135	CAACAACCTTGGAGTTCTTCGACCGCGAGTTGAGCGCATCGATGAGATTCTCCACGACGC	1328194
QY	601	CGGCATCAACAATTTGGCGCGAGGCCACCGATTCGAAACCCGCGAATACTAACCCATCAA	660
Db	1328195	CGGCATCAACAATTTGGCGCGAGGCCACCGATTCGAAACCCGCGAATACTAACCCATCAA	1328254
QY	661	CATCAGTTTGTATGGCCCAATGCGGTTCATCACTCAACTGCCACGACGAGTTGATCCAGGCGCA	720
Db	1328255	CATCAGTTTGTATGGCCCAATGCGGTTCATCACTCAACTGCCACGACGAGTTGATCCAGGCGCA	1328314
QY	721	CACCTTGGGGCTGGACAGCGGGCGGTGACAAATGCTGTCTGGCGCGAAACCCACGAGCGGAAA	780
Db	1328315	CACCTTGGGGCTGGACAGCGGGCGGTGACAAATGCTGTCTGGCGCGAAACCCACGAGCGGAAA	1328374
QY	781	CCAGATCAGGCTTCCGCGGAAACGGCGACGCGCGGAGAAATCCACGTCGCGGTGTCGCCGTA	840
Db	1328375	CCAGATCAGGCTTCCGCGGAAACGGCGACGCGCGGAGAAATCCACGTCGCGGTGTCGCCGTA	1328434
QY	841	TTTGGCGCCGACGCGCCGCGATAAACACAAACCGCTCCAAATACGATTCGGGGTTCACCA	900
Db	1328435	TTTGGCGCCGACGCGCCGCGATAAACACAAACCGCTCCAAATACGATTCGGGGTTCACCA	1328494
QY	901	GGTCAGACAGATTGCCATCAACATGGGCTTTACCCAAACCCGCTGCTTTATCGAGCTCAC	960
Db	1328495	GGTCAGACAGATTGCCATCAACATGGGCTTTACCCAAACCCGCTGCTTTATCGAGCTCAC	1328554
QY	961	CTCCACCCGACCCGGTTCGCGCTGTTCAGTGGGCCACCGCGCAACCGCCCAAGCGCTGTC	1020
Db	1328555	CTCCACCCGACCCGGTTCGCGCTGTTCAGTGGGCCACCGCGCAACCGCCCAAGCGCTGTC	1328614

[illegible]







Db 541 CGTGGCGGATGTGAAGTGTGAACTTGGAAACCATGCGCCACTTGGCCACTTGGCAATTCGAACCCC 600  
QY 1921 CTCATTGGGGGATGCTTACATGTTGATGGGAACCTAGATTGGGCTCGATGCCGTCTT 1980  
Db 601 CTCATTGGGGGATGCTTACATGTTGATGGGAACCTAGATTGGGCTCGATGCCGTCTT 660  
QY 1981 AGCTTTCGGTCCCAAGAGATGTCTTCAAGACCGTGACCTGGACGGGGCGCTCGATGGTCC 2040  
Db 661 AGCTTTCGGTCCCAAGAGATGTCTTCAAGACCGTGACCTGGACGGGGCGCTCGATGGTCC 720  
QY 2041 TGTGGGGCGAGGGCGGTATCCATTGTCCTCGTGGGGAGAGTGTGTTGGTGAAGCAATTCG 2100  
Db 721 TGTGGGGCGAGGGCGGTATCCATTGTCCTCGTGGGGAGAGTGTGTTGGTGAAGCAATTCG 780  
QY 2101 CGAGGCGCTTGGTGGGACTTCTCCGAAACCCAGAGCTGCTCCATGCTAAAGACAGG 2160  
Db 781 CGAGGCGCTTGGTGGGACTTCTCCGAAACCCAGAGCTGCTCCATGCTAAAGACAGG 840  
QY 2161 AGAAGTGTATCTCTCTCGATGAGATACCCATTTGACACACCGATGTATTGGCAACGATGGCG 2220  
Db 841 AGAAGTGTATCTCTCTCGATGAGATACCCATTTGACACACCGATGTATTGGCAACGATGGCG 900  
QY 2221 CTTGGATCTAGATCTCTAGCTAGACTCACAGAGCGGCTGCTGATGACAGCAATCGAGG 2280  
Db 901 CTTGGATCTAGATCTCTAGCTAGACTCACAGAGCGGCTGCTGATGACAGCAATCGAGG 960  
QY 2281 ATTGGGCGCTTAGTACTCTGAAAGGTTTCAG 2313  
Db 961 ATTGGGCGCTTAGTACTCTGAAAGGTTTCAG 993

RESULT 4  
US-09-738-626-3456  
; Sequence 3456, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 3456  
; LENGTH: 870  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3456  
  
Query Match 36.6%; Score 870; DB 9; Length 870;  
Best Local Similarity 100.0%; Fred. No. 6e-280;  
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1421 ATGAACCCCATTAACCTGGACACTTGTCTCTCAATCATTTGATGAAGGACGCTTCGAAGGC 1480  
Db 1 ATGAACCCCATTAACCTGGACACTTGTCTCTCAATCATTTGATGAAGGACGCTTCGAAGGC 60  
QY 1481 GCCTCTTAGCCCTTTCCATTTCCCTCGCGGTGAGTACAGCGGTAAAGCTCTCGAG 1540

Db 61 GCCTCTTAGCCCTTTCCATTTCCCTCGCGGTGAGTACAGCGGTAAAGCTCTCGAG 120  
QY 1541 CATCAGTGGGTGAGTGTGTTGATTCGCGCACCCAAACCGGCAAGCAACCGAAGCGGT 1600  
Db 121 CATCAGTGGGTGAGTGTGTTGATTCGCGCACCCAAACCGGCAAGCAACCGAAGCGGT 180  
QY 1601 GAAGTCTTGTGCAAGCAGCGCGGAATGTTGTCGCAAGCAGAACTTAAGCGCAA 1660  
Db 181 GAAGTCTTGTGCAAGCAGCGCGGAATGTTGTCGCAAGCAGAACTTAAGCGCAA 240  
QY 1661 CTATCTGCAAGCGCTTGTGAAATCCCGTTAAACCATCGCCATCAACGCGAGATTCGCTATCC 1720  
Db 241 CTATCTGCAAGCGCTTGTGAAATCCCGTTAAACCATCGCCATCAACGCGAGATTCGCTATCC 300  
QY 1721 ACATGTTTCTCCCGTGTTCACAGAGTGTCTTTCGGGTGAGAGCAACGCTCAGCGTG 1780  
Db 301 ACATGTTTCTCCCGTGTTCACAGAGTGTCTTTCGGGTGAGAGCAACGCTCAGCGTG 360  
QY 1781 CGCTTGAAGATGAAGCGCACATTTATCTTCTGCGGCGTGGAGATGTTTTAGAGCG 1840  
Db 361 CGCTTGAAGATGAAGCGCACATTTATCTTCTGCGGCGTGGAGATGTTTTAGAGCG 420  
QY 1841 GTAAACCGGTGAAGCTAATCCCGTGGCGGATGTGAAGTAGTAGAACTTGGAAACCATGCGC 1900  
Db 421 GTAAACCGGTGAAGCTAATCCCGTGGCGGATGTGAAGTAGTAGAACTTGGAAACCATGCGC 480  
QY 1901 CACTTGGCCATTCGAAACCCCTCATTTGGGGATGCTATCGTGTGAGGAACTAGAT 1960  
Db 481 CACTTGGCCATTCGAAACCCCTCATTTGGGGATGCTATCGTGTGAGGAACTAGAT 540  
QY 1961 TGGGCTGCGATGCGCTTTCAGCTTCCGTCCTCCAAAGATGTGCTTCAAGACCGTGACCTG 2020  
Db 541 TGGGCTGCGATGCGCTTTCAGCTTCCGTCCTCCAAAGATGTGCTTCAAGACCGTGACCTG 600  
QY 2021 GACGGCGCGTGTGATGCTTCTGTTGGGGCGCAGGCGGTATCCATTTGTCCTCCGCGGAA 2080  
Db 601 GACGGCGCGTGTGATGCTTCTGTTGGGGCGCAGGCGGTATCCATTTGTCCTCCGCGGAA 660  
QY 2081 GCTTTTGTGAGGCAATTCGCGAGCGCTTGGTTGGGGACTTCTCCGGAACCGAAGCT 2140  
Db 661 GCTTTTGTGAGGCAATTCGCGAGCGCTTGGTTGGGGACTTCTCTCCGGAACCGAAGCT 720  
QY 2141 GCTCCCATGCTAAAGCAGGAGAGTGTATCTCTCGATGAGATACCCATTTGACACACCG 2200  
Db 721 GCTCCCATGCTAAAGCAGGAGAGTGTATCTCTCGATGAGATACCCATTTGACACACCG 780  
QY 2201 ATGATTTGCAACGATGCGCGCTTGAATCTAGATCTCTAGCTAGACTCACAGAGCGCGTC 2260  
Db 781 ATGATTTGCAACGATGCGCGCTTGAATCTAGATCTCTAGCTAGACTCACAGAGCGCGTC 840  
QY 2261 GTTGATGCAAGCAATCGAGGATTCGGGCT 2290  
Db 841 GTTGATGCAAGCAATCGAGGATTCGGGCT 870

RESULT 5  
US-09-746-660A-51/c  
; Sequence 51, Application US/09746660A  
; Publication No. US20030049804A1  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schroder, Hartwig  
; APPLICANT: Zelder, Oskar  
; APPLICANT: Habernauer, Gregor  
; APPLICANT: Kim, Jun-Won  
; APPLICANT: Hwang, Byung-Joon  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING  
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS  
; FILE REFERENCE: BGI-121CP2  
; CURRENT APPLICATION NUMBER: US/09/746,660A

[illegible]

QY 1069 CTTGTTTGTCAATGGCGCTCTTTGCTGCGATGACGGCAAAACCAATACAGGTAAAGCATGCC 1128  
 Db 291 CTTGTTTGTCAATGGCGCTCTTTGCTGCGATGACGGCAAAACCAATACAGGTAAAGCATGCC 232  
 QY 1129 ACCCCAGCGCATTAATATCAGACGATCGCGCGGCAATGACAAAGATCAAGCCCAA 1188  
 Db 231 ACCCCAGCGCATTAATATCAGACGATCGCGCGGCAATGACAAAGATCAAGCCCAA 172  
 QY 1189 GTGTCGCGCGATGACAAAGAGCGTCAGAAATTAACACACGAGGAACCGCAATGAG 1248  
 Db 171 GTGTCGCGCGATGACAAAGAGCGTCAGAAATTAACACACGAGGAACCGCAATGAG 112  
 QY 1249 TCCCTTCGCGCTTAATTCCTTTGTTTAATCACCAGTACATCTCGCGTCCGATGGAAGTAA 1308  
 Db 111 TCCCTTCGCGCTTAATTCCTTTGTTTAATCACCAGTACATCTCGCGTCCGATGGAAGTAA 52  
 QY 1309 AGACTGCGCGCGCAAAAGAGCAGCTGTAAATGAAGATTTCCATGATCACCAT 1359  
 Db 51 AAGACTGCGCGCGCAAAAGAGCAGCTGTAAATGAAGATTTCCATGATCACCAT 1

RESULT 7

US-09-738-626-3455/c  
 ; Sequence 3455, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: Patent in ver. 3.0  
 ; SEQ ID NO 3455  
 ; LENGTH: 708  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-3455

Query Match 29.8%; Score 708; DB 9; Length 708;  
 Best Local Similarity 100.0%; Pred. No. 1e-225;  
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 652 ACCCATCAATCAGTTTGTATGGCCCAATGGCTCATCACTGCCACGACGATTGAT 711  
 Db 708 ACCCATCAATCAGTTTGTATGGCCCAATGGCTCATCACTGCCACGACGATTGAT 649  
 QY 712 CCAGCGCCACACCTTGGGCGCTGGACAGCGGCGTGACAAATGCTGCGCGCAAAACCCAC 771  
 Db 648 CCAGCGCCACACCTTGGGCGCTGGACAGCGGCGTGACAAATGCTGCGCGCAAAACCCAC 589  
 QY 772 CAGCGGGAAACAGATCAGCTTGGCGGAAACCGCGCAGCGCGGAAATCCACGTCGGGT 831  
 Db 588 CAGCGGGAAACAGATCAGCTTGGCGGAAACCGCGCAGCGCGGAAATCCACGTCGGGT 529  
 QY 832 GTCCGCGTATTCGCGCGACGCGCGGATTAACACAAACGCTCCAAATAGCATTCGG 891  
 Db 528 GTCCGCGTATTCGCGCGACGCGCGGATTAACACAAACGCTCCAAATAGCATTCGG 469

QY 892 GTTCAACAGGTGACGACGATTCATCAATGGGCTTTACCCAAACCGCGCTGTATTC 951  
 Db 468 GTTCAACAGGTGACGACGATTCATCAATGGGCTTTACCCAAACCGCGCTGTATTC 409  
 QY 952 GAGCTCTACCTCCACCGCACCGGTTGCGGCTGTCAAGTGGCCACCGCCGAAACCGCCAA 1011  
 Db 408 GAGCTCTACCTCCACCGCACCGGTTGCGGCTGTCAAGTGGCCACCGCCGAAACCGCCAA 349  
 QY 1012 AGGCTGTCTATCGCGCACCGGTTGCTGTTTCTCAATGATCTGTGGCGCTTCCACCTT 1071  
 Db 348 AGGCTGTCTATCGCGCACCGGTTGCTGTTTCTCAATGATCTGTGGCGCTTCCACCTT 289  
 QY 1072 GTTGTCTATCGCGCTTTCCTGCTGCCATGAGGCAAAACCATACAGGTAAAGTGCAC 1131  
 Db 288 GTTGTCTATCGCGCTTTCCTGCTGCCATGAGGCAAAACCATACAGGTAAAGTGCAC 229  
 QY 1132 CAGCGGCAATATCGAGACGATCGCGCGGATTTGACAAAAGATCAACGCGCCAAAGT 1191  
 Db 228 CAGCGGCAATATCGAGACGATCGCGCGGATTTGACAAAAGATCAACGCGCCAAAGT 169  
 QY 1192 GCGGCGGATGACAAAAGACGTCAGAAATTAACACACGAGGAACCGCAATGAGTCC 1251  
 Db 168 GCGGCGGATGACAAAAGACGTCAGAAATTAACACACGAGGAACCGCAATGAGTCC 109  
 QY 1252 TTCGCGCTTAATTCCTTGTGTTTAAATCACCAGTACATTTCTGCGTCCGATGGACAGTAAAG 1311  
 Db 108 TTCGCGCTTAATTCCTTGTGTTTAAATCACCAGTACATTTCTGCGTCCGATGGACAGTAAAG 49  
 QY 1312 ACTGCGCGCCCAAAAGACGACCTGTAAATGAGATTTCCATGATCACCAT 1359  
 Db 48 ACTGCGCGCCCAAAAGACGACCTGTAAATGAGATTTCCATGATCACCAT 1

RESULT 8

US-10-166-142-9/c  
 ; Sequence 9, Application US/10166142  
 ; Publication No. US20030124687A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GUNJI, YOSHIYA  
 ; APPLICANT: YASURDA, HISASHI  
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL  
 ; TITLE OF INVENTION: ASSIMILATING BACTERIUM  
 ; FILE REFERENCE: 223789US  
 ; CURRENT APPLICATION NUMBER: US/10/166,142  
 ; CURRENT FILING DATE: 2002-06-11  
 ; PRIOR APPLICATION NUMBER: JP 2001-177075  
 ; PRIOR FILING DATE: 2001-06-12  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 9  
 ; LENGTH: 712  
 ; TYPE: DNA  
 ; ORGANISM: Brevibacterium lactofermentum  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(375)  
 ; OTHER INFORMATION:  
 US-10-166-142-9

Query Match 29.4%; Score 696.8; DB 14; Length 712;  
 Best Local Similarity 99.6%; Pred. No. 5.7e-222;  
 Matches 709; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 QY 649 CTAACCCATCAATCAGTTTGTATGGCCCAATGGGTCATCACTGCCACGACGCTT 708  
 Db 712 CTAACCCATCAATCAGTTTGTATGGCCCAATGGGTCATCACTGCCACGACGCTT 653  
 QY 709 GATCCAGCGCCACACCTTGGGCGTGGACAGCGGCGTGACAAATGCTGTCGCGCGAAACC 768  
 Db 652 GATCCAGCGCCACACCTTGGGCGTGGACAGCGGCGTGACAAATGCTGTCGCGCGAAACC 593  
 QY 769 CACCAGCGGGAAACAGATCAGGCTTTCGCGGAAACGCGCGAGCGGGAATTCACCGTCC 828

Db 592 CACACGCGGAAACAGATAGGCTTCCGCGAGCGCGCGGAGAAATCCACCGTCC 533  
Qy 829 GGTGTGCGGATTTGCGCGCGGAGCGCGCGGATTAACAAACCGGTCGCAATTCGCAAT 888  
Db 532 GGTGTGCGGATTTGCGCGCGGAGCGCGCGGATTAACAAACCGGTCGCAATTCGCAAT 473  
Qy 889 CGGGTCAACAGGTGAGCAGATTTGCCATCAACATGGGCTTTACCCAAACCGGTCGCTT 948  
Db 472 CGGGTCAACAGGTGAGCAGATTTGCCATCAACATGGGCTTTACCCAAACCGGTCGCTT 413  
Qy 949 ATCGACGCTCACTCCACCGCGACCGCGGTTCGCGGTGTCAGTGGCGCACCGCGCAAC-CGC 1007  
Db 412 ATCGACGCTCACTCCACCGCGACCGCGGTTCGCGGTGTCAGTGGCGCACCGCGCAAC-CGC 353  
Qy 1008 CAAAGGCGGTGTCATCGGCGCAGCGGTGGTCTGTGTTCTTCAATGATCTGTGGCGCTTCCA 1067  
Db 352 CAAAGGCGGTGTCATCGGCGCAGCGGTGGTCTGTGTTCTTCAATGATCTGTGGCGCTTCCA 293  
Qy 1068 CTTTGTGTCATGGCGCTTTTCGCTGCCATGACGCGCAACCAATCAACAGGTAAAGCGATGC 1127  
Db 292 CTTTGTGTCATGGCGCTTTTCGCTGCCATGACGCGCAACCAATCAACAGGTAAAGCGATGC 233  
Qy 1128 CACCCAGCGCATTAATATCGACGACGATCGCGCGCGCATTTGGAACAAAGATCAACGCCCA 1187  
Db 232 CACCCAGCGCATTAATATCGACGACGATCGCGCGCGCATTTGGAACAAAGATCAACGCCCA 173  
Qy 1188 AGTGCGCGCGATGACAAAG 1247  
Db 172 AGTGCGCGCGATGACAAAG 113  
Qy 1248 GTCTTTCGCGCTTAATTCCTTTTAAATCAGCAGTACATTTGCGGTCCGATGACAGTA 1307  
Db 112 GTCTTTCGCGCTTAATTCCTTTTAAATCAGCAGTACATTTGCGGTCCGATGACAGTA 53  
Qy 1308 AAAGACTGCCCCCAAG 1359  
Db 52 AAAGACTGCCCCCAAG 1

RESULT 9  
US-09-738-626-3454  
; Sequence 3454, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738, 626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: Patent in ver. 3.0  
; SEQ ID NO 3454  
; LENGTH: 627  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3454

Query Match 26.4%; Score 627; DB 9; Length 627;  
Best Local Similarity 100.0%; Pred. No. 1.3e-198; Mismatches 0; Indels 0; Gaps 0;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 23 ATGTACGCATTCGCTGACATTTGCTGCTGGAAGGCTCTTTAGCTGGGTATTCTTCC 82  
Db 1 ATGTACGCATTCGCTGACATTTGCTGCTGGAAGGCTCTTTAGCTGGGTATTCTTCC 60  
Qy 83 TAGCGTCCAGAGCTCACAGCGGAGCGGCTGAGTTCATGCGGAGAGAGGCTGCCGCTT 142  
Db 61 TAGCGTCCAGAGCTCACAGCGGAGCGGCTGAGTTCATGCGGAGAGAGGCTGCCGCTT 120  
Qy 143 CTGATTCATCAGCAAGCTATTCATCATTAATGTTGGGTGAGGAAACCGCGGATGAC 202  
Db 121 CTGATTCATCAGCAAGCTATTCATCATTAATGTTGGGTGAGGAAACCGCGGATGAC 180  
Qy 203 GGTGAGAACTTTGTGAGTCAAGTCCCAACAAATGCTTTGGGCTCAATTCCTTCCACA 262  
Db 181 GGTGAGAACTTTGTGAGTCAAGTCCCAACAAATGCTTTGGGCTCAATTCCTTCCACA 240  
Qy 263 CTTGCGCAGGCGCTCTCCGAGCAATATCTCGATGGAATTCAGAGGCTTCCCGGCGC 322  
Db 241 CTTGCGCAGGCGCTCTCCGAGCAATATCTCGATGGAATTCAGAGGCTTCCCGGCGC 300  
Qy 323 AGCCAGGTAAGTCCCTGCTCTGAGGCGCATTTGAAACGTAATTAATGATGTCGCGC 382  
Db 301 AGCCAGGTAAGTCCCTGCTCTGAGGCGCATTTGAAACGTAATTAATGATGTCGCGC 360  
Qy 383 AAGCTCAATGATCGCGCGGAGGAGCGCGGCGATCACTTCCGAGATGCGCTTGCATGG 442  
Db 361 AAGCTCAATGATCGCGCGGAGGAGCGCGGCGATCACTTCCGAGATGCGCTTGCATGG 420  
Qy 443 GTGCTGCGGAGGAGAGAGTACGCGCGGATACCGTACGAGTGCATTTGATGTCGCT 502  
Db 421 GTGCTGCGGAGGAGAGAGTACGCGCGGATACCGTACGAGTGCATTTGATGTCGCT 480  
Qy 503 TCGTCAGTTGAGCAGCTGGAACAAACCGCTTGAATTCATCAACACTTTGAGTTTCGAC 562  
Db 481 TCGTCAGTTGAGCAGCTGGAACAAACCGCTTGAATTCATCAACACTTTGAGTTTCGAC 540  
Qy 563 GCGAGTTGAGGCGATGATGAGATTTCCACAGCGCGGATCAACATTTGGGCGAG 622  
Db 541 GCGAGTTGAGGCGATGATGAGATTTCCACAGCGCGGATCAACATTTGGGCGAG 600  
Qy 623 GCCACCGATTCCAAAACCGCGGAAAC 649  
Db 601 GCCACCGATTCCAAAACCGCGGAAAC 627

RESULT 10  
US-10-627-476-669  
; Sequence 669, Application US/10627476  
; Publication No. US20040030116A1  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Mark  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schoder, Hartwig  
; APPLICANT: Zelder, Oskar  
; APPLICANT: Haberhauser, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
; INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE  
; TITLE OF INVENTION: TRANSPORT  
; FILE REFERENCE: BGI-125CPCN  
; CURRENT APPLICATION NUMBER: US/10/627,476  
; CURRENT FILING DATE: 2003-07-25  
; PRIOR APPLICATION NUMBER: 09/602,787  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: USSN 60/141031  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: DE 19931454.3  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931478.0  
; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19931563.9  
 ; PRIOR FILING DATE: 1999-07-08  
 ; PRIOR APPLICATION NUMBER: DE 19932122.1  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932124.8  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932125.6  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932128.0  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932180.9  
 ; PRIOR FILING DATE: 1999-07-09  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 678  
 ; SEQ ID NO 669  
 ; LENGTH: 993  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (101)..(970)  
 ; OTHER INFORMATION: RKN03164  
 US-10-627-476-669

Query Match 20.4%; Score 485.2; DB 12; Length 993;  
 Best Local Similarity 99.0%; Pred. No. 5.8e-151;  
 Matches 499; Conservative 0; Mismatches 3; Indels 2; Gaps 1;  
 QY 1 AGTACTCTTTGGAGAACCATGTAGCATTTGGTGCATTTGTCGCTTGGAAAGGC 60  
 DB 490 AGTACTCTTTGGAGAACCATGTAGCATTTGGTGCATTTGTCGCTTGGAAAGGC 549  
 QY 61 TCTTTACGTGGTATTTCTTCTACGCTCAGAGCTCAGCGAGCGGCTGAGTTTCA 120  
 DB 550 TCTTTACGTGGTATTTCTTCTACGCTCAGAGCTCAGCGAGCGGCTGAGTTTCA 609  
 QY 121 GCGGAGAGGGCTGCGCGCTTCTGATTCATCAGCAAGCTATTCATCAATTAATCGTTG 180  
 DB 610 GCGGAGAGGGCTGCGCGCTTCTGATTCATCAGCAAGCTATTCATCAATTAATCGTTG 669  
 QY 181 GGTGAGAGAACCGGGGATGACGGTGGAGACTTTGGCAGTCAGCTGCCAACAATGCTT 240  
 DB 670 GGTGAGAGAACCGGGGATGACGGTGGAGACTTTGGCAGTCAGCTGCCAACAATGCTT 729  
 QY 241 TGGCGTCAATTCCTTTCTACCACTTGGCAGGCGCTGCTCAGCAACAATTCATCGATGG 300  
 DB 730 TGGCGTCAATTCCTTTCTACCACTTGGCAGGCGCTGCTCAGCAACAATTCATCGATGG 789  
 QY 301 AATTCAGAGGGTTCGCGCGCAGCCAGGGTAAAGTCCCTGTCTGAGGCAATGTTGAAGT 360  
 DB 790 AATTCAGAGGGTTCGCGCGCAGCCAGGGTAAAGTCCCTGTCTGAGGCAATGTTGAAGT 849  
 QY 361 GAACATATTCATATGTCGCAAGCTCAATGATCGCCAGCAAGAGCGGCGAGTCACT 420  
 DB 850 GAACATATTCATATGTCGCAAGCTCAATGATCGCCAGCAAGAGCGGCGAGTCACT 909  
 QY 421 TGGCAGATGCGGCTTGCATGGGTGCTGCGCAGCAAGAGAGTACCGGCGC--GGATACC 478  
 DB 910 TGGCAGATGCGGCTTGCATGGGTGCTGCGCAGCAAGAGAGTACCGGCGCGGATACC 969  
 QY 479 GTGACAGTGCATGATGGTGTCT 502  
 DB 970 GTGACAGTGCATGATGGTGTCT 993

RESULT 11  
 US-10-282-122A-17521  
 ; Sequence 17521, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SEQ ID NO 17521  
 ; SOFTWARE: PatentIn version 3.1  
 ; LENGTH: 1095  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium diptheriae  
 US-10-282-122A-17521

Query Match 12.9%; Score 305.8; DB 12; Length 1095;  
 Best Local Similarity 67.1%; Pred. No. 8e-91;  
 Matches 433; Conservative 0; Mismatches 212; Indels 0; Gaps 0;  
 QY 1 AGTACTCTTTGGAAGAACCATGTAGCATTTGGCTGACATTTGCGCTTGAAGGC 60  
 DB 450 AGTACTCTTTGGAAGAACCATGTAGCATTTGGCTGACATTTGCGCTTGAAGGC 509  
 QY 61 TCTTTACGTGGTATTTCTTCTACGCTCAGAGCTCAGCGAGGCGGCTGAGTTTCA 120  
 DB 510 TCGCTATGTGGGATTTCTTCTTATGGACCTGAACTAACCGCAGAGATCTCGAATTTCT 569  
 QY 121 GCGGAGAGGGCTGCGCGCTTCTGATTCATCAGCAAGCTATTCATCAATTAATCGTTG 180  
 DB 570 AGCTGTGAGGGCTGCGCGCTTCTGATTCATCAGCAAGCTATTCATCAATTAATCGTTG 629  
 QY 181 GGTGAGAGAACCGGGCGATGACGGTGGAGACTTTGGCAGTCAGCTGCCAACAATGCTT 240  
 DB 630 GGTGAGAGAACCGGGCGATGACGGTGGAGACTTTGGCAGTCAGCTGCCAACAATGCTT 689  
 QY 241 TGGCGTCAATTCCTTCTACCACTTGGCAGGCGCTGCTCAGCAACAATTCATCGATGG 300  
 DB 690 GGTGTATTTGCAATTTTCGCCACTTGCACAGGGGCTTTTAACTGACAGGTATCTCGAGG 749  
 QY 301 AATTCAGAGGGTTCGCGCGCAGCAAGCTGAGTAAAGTCCCTGTCTGAGGCGCATGTTGAAGT 360  
 DB 750 CGTGCCAGTGGATTCACGCGCAGCAGCAGGGAAGTCTTTGGGCAAGAGATGCTCAAGC 809  
 QY 361 GAACATATTCATATGTCGCAAGCTCAATGATCGCCAGCAAGAGCGGCGAGTCACT 420  
 DB 810 TAAAAATCTTGACATGGTGGTGCATGAATGATTTAGGGTTTACGACGTGGCCAAAGCTT 869

QY 421 TGGCAGATGGCGCTTGGATGGTGGGAGCAGAGAGTACGGCGGATACCT 480  
Db  
870 GGCACAGCTGGCGTATTGGTGGGTGCTGGGGAAACAGAGGATATTGGGGCCAAACGGT 929  
QY 481 GACAGTGCATTGATTGGTGTCTTCGTCAGTTGAGCAGCTGGACCAAGCCCTTGATCACT 540  
Db 930 AACAGTGGGTGATCGAGGATCGTCTGGTAGCGCATTTGGATCAAAATCTGGAGCTCT 989  
QY 541 CAACAATCTGGAGTTTCTGACCGCAGTGGAGGCGGATCGATGAGATTTCCCAAGACGC 600  
Db 990 TAAATATCTGATGTCAGCGTTGAGGAACGCAATATATGACATCTGTTGCCAAGGACGC 1049  
QY 601 CGGCATCAACATTTGGCGAGCGCCACCGATTTCCAAAACCGCGA 645  
Db 1050 TGGAAATTAATTTGGCGGCGTGTACTGCTCGAAGGTTACGA 1094

RESULT 12

US-10-282-122A-17523  
; Sequence 17523, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17523  
; LENGTH: 879  
; TYPE: DNA  
; ORGANISM: *Corynebacterium diptheriae*  
US-10-282-122A-17523  
Query Match 12.8%; Score 302.8; DB 12; Length 879;  
Best Local Similarity 59.2%; Pred. No. 7e-90;  
Matches 517; Conservative 0; Mismatches 357; Indels 0; Gaps 0;  
QY 1421 ATGAACCCCATTCACACTGGACACTTTGCTCTCATCATCATGATGAGCGAGCTTCGAAGGC 1480

Db 1 ATGAATCCGCTTCATTAGAAACCTTGTCTCGCAATATTCGACGAGCGAGCTTCGAAGAT 60  
QY 1481 GCCTCCTTAGCCTTTCCATTTCCCTCCCTCGCGGTGAGTACGCGGTAAAGCTCTCGAG 1540  
Db 61 CGCGCTTTTGCAATTAGAAATTTCTCTTCAGCAGTAAGTACAGAGGATTAAGCGCTTGAG 120  
QY 1541 CATCAGTGGTGGTGTGTGTGTATCGCCACCCCAACCGGCCAAAGCAACGAGCGGCT 1600  
Db 121 CGGAGAGACTGGTGTGAGTGTGTGTATCGCCGTACGTCAAGTCACTGCACTGTATGCGGC 180  
QY 1601 GAAGTCCCTTGTGCAAGCAGCGCGGAAATGTGTGTGTGCAAGCAGAAATTAAGCGCAA 1660  
Db 181 GAGGTACTGTGTCAATCGGCACGAGTATGGCACTTTCGACGCGCCGAACAAAGCTCAG 240  
QY 1661 CTATCTGGAGCGCTGTGTGAAATCCCGTTAAACCATCGCATCAAGCAGATTCGCTATCC 1720  
Db 241 TTGCGGGGAAGAAATTGACCGGATTTCCATTTGTCTGTAGCTGTCAACGAGATTTCTTGCC 300  
QY 1721 ACATGGTTTCTCCGCTGTTCACGAGGTAGCTTCTTGGGGTGGAGCAACGCTCACGCTG 1780  
Db 301 ACATGGTTTCTCCGCGGTGTATGCGCAAGTGTGGCGGAGATGATTAACGCCACATTCATG 360  
QY 1781 CGCTTGGAGATGAAGCGCACACATTTATCTTGTGTGGCGGTGGAGATGTTTTAGAGCG 1840  
Db 361 GGTATCGAGGAGAGTGTGCAATTCGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420  
QY 1841 GTAAACCGGTGAGCTTAATCCCGTGGCGGAGTGTGAAGTAGTAGAACTTGAACCATTCGC 1900  
Db 421 GTAAACCGGGGAAGACACACTGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480  
QY 1901 CACTTGGCCTATTGCAACCCCTCATTTGCGGGATGCTCATGTTGATGGGAAACTAGAT 1960  
Db 481 TATTTTGCAGTGTGCAAGTTCGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540  
QY 1961 TGGGCTCGAGTGGCGCTCTTAGCTTGGTGTCCAAAGATGTGTTCAAGACGAGCTG 2020  
Db 541 TGGGAGCGGATGCCAGTATTGCGATTTGGCCCTCGGGATGGGCTTCAGACCGTGATTT 600  
QY 2021 GACGGGCGCTGAGT 2080  
Db 601 AAGCGTGTCTTGTGTGAAGTTCGCGGAGGCGGACGTGTATGAGATTCCTTCATCGGAG 660  
QY 2081 GGTGTGTGTGAGGCAATTCGCGGAGGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2140  
Db 661 GCATTCATGGAAGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720  
QY 2141 GCTCCCATGTAAAGCAGGAGAGTGTATCTCTCTCGATGAGTACCCATTGACACACCG 2200  
Db 721 TTACCGCTTTTGGATTTCTGT 780  
QY 2201 ATGTATTGGCAACGATGGCGCTCGAATCTAGATCTCTAGCTAGACTCACAGACGCGTC 2260  
Db 781 CTGTATTGGCAGCGATGGCGCTTGAATCTCCAGCCCTAGAGCGGTTGACTCAAGCGTA 840  
QY 2261 GTGTATGCACAATCAGGAGATTCGGCCCTTAGT 2294  
Db 841 GTGGATGCAAGCGCTTCAGGAGACTACATCCACTGT 874

RESULT 13

US-10-369-493-35374  
; Sequence 35374, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 35374  
 ; LENGTH: 1026  
 ; TYPE: DNA  
 ; ORGANISM: Agrobacterium tumefaciens  
 US-10-369-493-35374

Query Match 6.9%; Score 164.4; DB 15; Length 1026;  
 Best Local Similarity 57.2%; Pred. No. 1.8e-43;  
 Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;

QY	3	ATACCTCTTTGGAGAAACATGACGATGCGGCGCTGGACAGATCGTGGCTCCGCGCAAGGCGC	62
DB	428	ACACACCGCTTGGAGAAACCTCGGCGCTGGACAGATCGTGGCTCCGCGCAAGGCGC	487
QY	63	TTTACGCTGGGTATTTCTTCTACGCTCCAGAGCTCAGCGGAGCGGCTGAGTTTCATGG	122
DB	488	TCTATGTGCGCATCTCTCTACACTGAGAGCGCACCGCGAGCGCGCTATCTCTGA	547
QY	123	CGGAGAGGGCTGCGCGCTCTGATTCATCAGCAAGCTATTCATCAATTAATCGTTGG	182
DB	548	AGGATCTCGGCAAGCGCTGATCATCCACAGCGGAGCTATTCGATGATCAACCGCTGA	607
QY	183	TGGAGAGCGGCGGATGAGCTGAGAACTTTGTCAGTCACTGCCAATGCTCTTG	242
DB	608	TCGAGAGAGCGGCTTTGTCG-----ATACGCTGGAAGAACTGGGTATCG	652
QY	243	GGCTCAATGCTTTCTACCACTTGGCAGGCGCTGCTCAGGACAAATATCTCGATGAA	302
DB	653	GCTCATGCTTTTTCGCGCTGGCGAGGGCATGCTGACGACGAATATCTGGCGGTG	712
QY	303	TTCCAGAGGGTTCCGCGCAGCAAGGTAAGTCCCTGTCTGAGGGCATGTTGAACGTGA	362
DB	713	TGCGGATGGCAGCGCTGCTCAGAGCAAGTCACTCAACCGCGCTTCTCAACGAGC	772
QY	363	ACAATATGATGCTTCCAGAGCTCAATGACATGCCAGAGCGGCGGCTGCTCTTG	422
DB	773	GCAATGCGAAACATCCGCGCGCTGAAACAGCATTCGCGAGCGGCTGGCAGAGCTGG	832
QY	423	CGCAGATGGCGCTTGCATGGGTGCTGCGGAGCAAGAGAGTACGGCGCGATACCGTGA	482
DB	833	CGCAGATGGCAATTCCTGCGGTTCTGCGCGCGCG-----CATTA	874
QY	483	CCAGTGCATGATGCTGCTGCTGAGCAGCTGGAACAACAGCTTGAATCACTCA	542
DB	875	CCTCAGCATGATGGCGCAAGCGGTGCGAAACAGGTCGAGGACTGCGTGAAGCACTCG	934
QY	543	ACAACCTTGGAGTTTCTGACGCGGAGTTGGAGCGGATGATGAGATTTCCACGAGCGG	602
DB	935	ATAATGCGGAGTCTCTACCGAGGAGCTGGCGGAATGACCGTTACGCCAAGGATGCG	994
QY	603	GCATCAACATTTGGCGGAAGGCCACCGA	630
DB	995	ATATCAACCTCTGGCGAAATCTGCCGA	1022

RESULT 14  
 US-10-369-493-38826  
 ; Sequence 38826, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 38826  
 ; LENGTH: 1026  
 ; TYPE: DNA  
 ; ORGANISM: Agrobacterium tumefaciens  
 US-10-369-493-38826

Query Match 6.9%; Score 164.4; DB 15; Length 1026;  
 Best Local Similarity 57.2%; Pred. No. 1.8e-43;  
 Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;

QY	3	ATACCTCTTTGGAGAAACATGACGATGCGGCGCTGGACAGATCGTGGCTCCGCGCAAGGCGC	62
DB	428	ACACACCGCTTGGAGAAACCTCGGCGCTGGACAGATCGTGGCTCCGCGCAAGGCGC	487
QY	63	TTTACGCTGGGTATTTCTTCTACGCTCCAGAGCTCAGCGGAGCGGCTGAGTTTCATGG	122
DB	488	TCTATGTGCGCATCTCTCTACACTGAGAGCGCACCGCGAGCGCGCTATCTCTGA	547
QY	123	CGGAGAGGGCTGCGCGCTTCTGATTCATCAGCAAGCTATTCATCAATTAATCGTTGG	182
DB	548	AGGATCTCGGCAAGCGCTGATCATCCACAGCGGAGCTATTCGATGATCAACCGCTGA	607
QY	183	TGGAGAGCGGCGGATGAGCTGAGAACTTTGTCAGTCACTGCCAATGCTCTTG	242
DB	608	TCGAGAGAGCGGCTTTGTCG-----ATACGCTGGAAGAACTGGGTATCG	652
QY	243	GGCTCAATGCTTTCTACCACTTGGCAGGCGCTGCTCAGGACAAATATCTCGATGAA	302
DB	653	GCTCATGCTTTTTCGCGCTGGCGAGGGCATGCTGACGACGAATATCTGGCGGTG	712
QY	303	TTCCAGAGGGTTCCGCGCAGCAAGGTAAGTCCCTGTCTGAGGGCATGTTGAACGTGA	362
DB	713	TGCGGATGGCAGCGCTGCTCAGAGCAAGTCACTCAACCGCGCTTCTCAACGAGC	772
QY	363	ACAATATGATGCTTCCAGAGCTCAATGACATGCCAGAGCGGCGGCTGCTCTTG	422
DB	773	GCAATGCGAAACATCCGCGCGCTGAAACAGCATTCGCGAGCGGCTGGCAGAGCTGG	832
QY	423	CGCAGATGGCGCTTGCATGGGTGCTGCGGAGCAAGAGAGTACGGCGCGATACCGTGA	482
DB	833	CGCAGATGGCAATTCCTGCGGTTCTGCGCGCGCG-----CATTA	874
QY	483	CCAGTGCATGATGCTGCTGCTGAGCAGCTGGAACAACAGCTTGAATCACTCA	542
DB	875	CCTCAGCATGATGGCGCAAGCGGTGCGAAACAGGTCGAGGACTGCGTGAAGCACTCG	934
QY	543	ACAACCTTGGAGTTTCTGACGCGGAGTTGGAGCGGATGATGAGATTTCCACGAGCGG	602
DB	935	ATAATGCGGAGTCTCTACCGAGGAGCTGGCGGAATGACCGTTACGCCAAGGATGCG	994
QY	603	GCATCAACATTTGGCGGAAGGCCACCGA	630
DB	995	ATATCAACCTCTGGCGAAATCTGCCGA	1022

RESULT 15  
 US-10-369-493-38124  
 ; Sequence 38124, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 38124  
 ; LENGTH: 1028  
 ; TYPE: DNA  
 ; ORGANISM: Agrobacterium tumefaciens  
 US-10-369-493-38124

Query Match 6.9%; Score 164.4; DB 15; Length 1028;  
 Best Local Similarity 57.2%; Pred. No. 1.8e-43;  
 Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;

Qy	3	ATACCTCTTTGGAAGAAACCATGTACGCAATTCGCTGACATTTGTTGGTCTGGAAGGCTC	62
Db	427	ACACACGGCTTCAGGAACCTCGCGCGCTGGACCATGCTGGCTCCGGCAAGGCGC	486
Qy	63	TTTACGTGGGTATTTCTCTACGTTCAGAGCTCACAGCGAGGGGCTGAGTTTATGS	122
Db	487	TCATGTTCGGCATCTCTCTTACAACTCGAAGCGCACCCCGGAGGCGCGCTATCTCTGA	546
Qy	123	CGAGGAGGGCTGCGCGCTTCTGATTCATCAGCAAGCTATTCATTAATTCGTTGGG	182
Db	547	AGGATCTCGGACGCGCTGCAATCTCACCGCGAGCTATTCGATGATCAACCGTGA	606
Qy	183	TGAGGAACCGGCGGATGACGGTGAAGAACTTGTTCAGTCAGCTGCCAAATGTTTG	242
Db	607	TCGAGGAAGACGGTCTTGTG-----ATACGCTGGAAGAACTGGGTATCG	651
Qy	243	CGGTCTATGCTTCTACCACTTCGCGAGGCGCTGCTCACGACAAATATCTCGATGGA	302
Db	652	GCTCCATCGTCTTTTGGCGCTGGCGAGGCGATGCTGACGACGAAATATCTGGGCGGTG	711
Qy	303	TTCCAGAGGGTTCGCGCGCCAGCCAGGCTAAGTCCCTGCTCTGAGGGCATGTTGAACGTGA	362
Db	712	TGCCGGATGCGACCGCTGCTCACAGAGCAAGTCACTCAACCGGCTTCTCTCAACGAGC	771
Qy	363	ACAATATTGATGTTGTCGCAAGCTCAATGACATCGCCAGGAAACCGGCGAGTCACTTG	422
Db	772	GCAATGTCGAAACATCCGCGCGCTGAACAGCAATTCGCGAGCGGCTGGCCAGAGCGTGG	831
Qy	423	CGCAGATGCGGCTTGCATGGGTGCTGCGGACAGGAGAGTACGGCGGCGATACCGTGA	482
Db	832	CGCAGATGCGCAATTCGCTGGGTCTGCGGCGGCGCG-----CATTA	873
Qy	483	CCAGTGCAATTGATTGCTGCTTCAGTTGAGCAGCTGGCAACAGCCTTGATTCACTCA	542
Db	874	CCTCAGCAATTGATTGGCGCAAGCCGTGTCGACAGGTCGAGGACTGGGTGAAGCACTCG	933
Qy	543	ACAATTTGGAGTTTCTGACGCGAGTTGGAGCGATCGATGAGATTCCCAACGACCG	602
Db	934	ATAATGCCGAGTTCTCTACCGAGGAGCTGGCCGAAATCGACCGTTACGCCAAGGATCGG	993
Qy	603	GCATCAACATTTGGCGAAGCCACCGA	630
Db	994	ATATCAACCTCTGGGCAAAATCTGCCGA	1021

Search completed: March 17, 2004, 09:47:27  
 Job time : 866.071 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:43:00 ; Search time 6528.5 Seconds  
(without alignments)  
10858.975 Million cell updates/sec

Title: US-09-105-117k-3  
Perfect score: 2374

Sequence: 1 agatactccttggagaagaa.....gtaacaccttcagcaaatgg 2374

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pin:\*
- 20: em\_gss\_vit:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	84.6	3.6	657	13 BQ155035	BQ155035 NF075E081
C 2	73.2	3.1	3237	28 BH770958	BH770958 LLMGtag68
3	72	3.0	860	28 AF075981	AF075981 AF075981
4	60.4	2.5	830	29 CNS01MQH	AL151258 Anopheles

C	5	56.8	2.4	1034	28	BZ554294	BZ554294
	6	54.2	2.3	897	28	BZ560550	BZ560550
	7	51.4	2.2	1620	28	BZ568946	BZ568946
	8	48.8	2.1	1101	29	CNS017SY	CNS017SY
	9	48.6	2.0	833	28	BZ556963	BZ556963
C	10	48.6	2.0	925	29	CNS0091P	CNS0091P
C	11	47.6	2.0	852	13	BX393687	BX393687
	12	46	1.9	614	14	CA830247	CA830247
C	13	46	1.9	720	14	CD881613	CD881613
	14	44.4	1.9	620	14	CA829022	CA829022
	15	44.4	1.9	644	13	BQ704105	BQ704105
C	16	44.4	1.9	664	29	CNS03K4J	CNS03K4J
C	17	44.4	1.9	1201	13	BX381961	BX381961
	18	43.6	1.8	721	28	BZ562255	BZ562255
	19	43.2	1.8	1395	28	BZ574094	BZ574094
C	20	42.8	1.8	532	10	AW927407	AW927407
C	21	42.8	1.8	538	14	CD861975	CD861975
	22	42.4	1.8	699	13	BX424825	BX424825
C	23	42.2	1.8	885	13	BX425603	BX425603
	24	42	1.8	457	13	BY246241	BY246241
	25	42	1.8	1201	13	BX381961	BX381961
	26	41.8	1.8	947	29	CNS077MV	CNS077MV
	27	41.8	1.8	1083	29	CNS079WN	CNS079WN
C	28	41.6	1.8	432	14	CB639421	CB639421
	29	41.6	1.8	558	12	BM712036	BM712036
	30	41.6	1.8	588	13	BU220119	BU220119
C	31	41.6	1.8	772	14	CB668778	CB668778
C	32	41.6	1.8	803	14	CB669972	CB669972
C	33	41.6	1.8	813	14	CB684812	CB684812
C	34	41.6	1.8	1201	13	BX356664	BX356664
C	35	41.6	1.8	1201	13	BX361080	BX361080
C	36	41.2	1.7	910	29	CNS006QN	CNS006QN
	37	41	1.7	617	12	BM729055	BM729055
	38	41	1.7	770	10	BE898495	BE898495
	39	41	1.7	861	10	BE740880	BE740880
	40	40.8	1.7	605	13	CA127255	CA127255
	41	40.8	1.7	614	13	CA141725	CA141725
	42	40.8	1.7	619	14	CA192339	CA192339
	43	40.8	1.7	619	14	CA238619	CA238619
	44	40.8	1.7	950	11	AV104298	AV104298
C	45	40.6	1.7	436	28	AQ405681	AQ405681

## ALIGNMENTS

RESULT 1  
BQ155035/c 657 bp mRNA linear EST 24-APR-2002  
LOCUS NF075E081R1F1067 Irradiated Medicago truncatula cDNA clone  
DEFINITION NF075E081R 5', mRNA sequence.

ACCESSION BQ155035  
VERSION BQ155035.1 GI:20292094  
KEYWORDS EST:  
SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE 1 (bases 1 to 657)  
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,  
Flores, H.R., Imman, J.T., Weller, J.W. and May, G.D.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula irradiated library

JOURNAL Unpublished (2001)  
COMMENT Contact: May GD

Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org

Insert Length: 657 Std Error: 0.00  
 Plate: 075 row: E column: 08  
 Seq primer: TCACAGGAACAGCTATGAC.

# FEATURES

Location/Qualifiers  
 1. .657  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3880"  
 /clone="NF075E081R"  
 /tissue\_type="seedlings"  
 /dev\_stage="seedling"  
 /clone\_lib="Irradiated"  
 /note="Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from poly(A)+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

# ORIGIN

Query Match 3.6%; Score 84.6; DB 13; Length 657;  
 Best Local Similarity 48.2%; Pred. No. 3.4e-11;  
 Matches 237; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 1420 CATGAACCCCATTCACCTGACACCTTGTCTCTCAATCATGATGAAGCGAGTTCGAAAG 1479  
 DB 605 CCNCGACTACAGACCATACAGCACTGGATCGGTGATACGTGAACGAGATTGAGCG 546

QY 1480 CGCTCTTACGCTTTCCATTTCCCTCGGCGGTGAGTCAGCGGTTAAAGCTCTCGA 1539  
 DB 545 CGCGGACAAAAGCTGTGCAATACACAAATCAGCGCTCTCACAGGCGCATTAAGCAACTGGA 486

QY 1540 GCATCACTGGGTGAGTGTGGTATCGGCACCCCAACCGCCGACGACGACGAGCGGG 1599  
 DB 485 AATATGTTGCGGCGAGCGCGTGTGGTGGTACCGTACCGCGCGCCGACGCAACAGG 426

QY 1600 TGAAGTCTCTGTGACAGCGCGGAAAATGTTGTGTGCAAGCAGAACTAAAGCGCA 1659  
 DB 425 GCNAAAATCTGTGCACTGCTGCGCAGGTGGAGTGTCTGGNAGAGAGTGGCTGGCGA 366

QY 1660 ACTATCTGGAGCGCTTGCTGAATAATCCGTTAAACATCGCCATCAACGAGATTCGCTATC 1719  
 DB 365 TGAACAAACCGGTTGCACTCCGCTGCTTTCACTGGCGGTCAACGCCGACAGTCTGGC 306

QY 1720 CACATGTTTCTCCCGTGTTCACAGAGTAGCTTCTTGGGTGGAGCAACGCTACCGT 1779  
 DB 305 GAGTGTGTGTTCTCTGCACTGGCTCTGTTGGCTGATTCGCGCTATCCGCTCACTT 246

QY 1780 GCGCTTGAAGATGAAGCGCACATTTATCTTGTCTGCGCGGTGGAGATGTTTATAGAGC 1839  
 DB 245 GCAGGTAGAGATGAACCGCACTCAGGAACGCTTCGCCCGCGGAGTGTGGCGGC 186

QY 1840 GGTAAACCGCTGAAGCTATCCCGTGGCGGATGTGAAGTAGTAGACTTGGAAACCAATGCG 1899  
 DB 185 GGTGAGTATTCAACATCAGCGCGTCCGAGTGTGTCTGTGCGATAAATCTGGTCCGCTGA 126

QY 1900 CCATTGGCCAT 1911  
 DB 125 CTATCTGTTCTGT 114

RESULT 2  
 BH770958/LOCUS 3237 bp DNA linear GSS 01-MAY-2002  
 DEFINITION LLMGtag686 MG1363 Random Sequence Tag Library Lactococcus lactis

# ACCESSION VERSION KEYWORDS SOURCE ORGANISM

# REFERENCE AUTHORS TITLE JOURNAL COMMENT

subsp. cremoris genomic, genomic survey sequence.  
 BH770958  
 BH770958.1 GI:20373915  
 GSS.  
 Lactococcus lactis subsp. cremoris  
 Lactococcus lactis subsp. cremoris  
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 Lactococcus.  
 1 (bases 1 to 3237)  
 Bolotin,A., Ehrlich,S.D. and Sorokin,A.  
 Studies of genomes of dairy bacteria Lactococcus lactis  
 Sci. Aliments (2002) In press  
 Contact: Sorokin A  
 Genetique Microbienne  
 INRA  
 CSJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France  
 Tel: 33 1 34 65 25 16  
 Fax: 33 1 34 65 25 21  
 Email: sorokine@jouy.inra.fr  
 best homologue in strain IL1403 is nadR (98%)  
 Class: shotgun  
 High quality sequence start: 30  
 High quality sequence stop: 3207.  
 Location/Qualifiers  
 1. 3237

# FEATURES source

/organism="Lactococcus lactis subsp. cremoris"  
 /mol\_type="genomic DNA"  
 /strain="MG1363"  
 /sub\_species="cremoris"  
 /db\_xref="taxon:1359"  
 /clone\_lib="MG1363 Random Sequence Tag Library"  
 /note="Vector: PSGM02; Site 1: SmaI; Library of chromosomal fragments of L.lactis strain MG1363 was prepared by partial AluI digestion or by sonication."

# ORIGIN

Query Match 3.1%; Score 73.2; DB 28; Length 3237;  
 Best Local Similarity 49.8%; Pred. No. 9.5e-08;  
 Matches 287; Conservative 0; Mismatches 263; Indels 26; Gaps 3;

QY 1 AGATCTCTCTTGAAGAAACCATGTACGATTCGCTGACATTTGTCGTCGAAAGC 60  
 DB 2511 AATATCCCTTATTCAGAAACCATGCGCGCTTTGAAACTGCTCGATAGTGGAAAGC 2452

QY 61 TCTTACGTGGGTATTTCTTCTTACGGTCCAGAGCTCACAGCGAGCGGTGAGTTCA 120  
 DB 2451 ACTTTATGTGGTGTGTCTAATTTATTTCTGCCAAGAAACCGAAGCTGCGGTCTTAGCGC 2392

QY 121 GCGGAGAGGGCTGCGCGCTCTGATTCATCAGCCAGCTATTCATCATTAATCGTTG 180  
 DB 2391 TGAAGATTAGGCTTTAAACTCTTGATTCATCAGCCTCGTTATCAATGTTAGATCGATG 2332

QY 181 GGTGAGGAACCGGCGCATGACGGTGAAGACTTGTTCAGTCACTGCCAACAATGGTCT 240  
 DB 2331 GATTGAGATGATTTACAGAACTTTGACAGAGGGG-----GGAT 2289

QY 241 TGGCGTCAATGCTTCTCAACACTTSCGAGCGCTGCTCAAGCAAAATATCTCGATGG 300  
 DB 2288 AGGAACAATTTGCTTTAAGCTTTGATCAAGGACTCTTAAACAGGAAATATTTGATGG 2229

QY 301 AATTCAGAGGGTTC-----CGCGCCAGCGAGGTAGTCCCTGCTGAGGGCATGTT 354  
 DB 2228 AATTCCTGAAATTCGCGAATCGGTGATCTCTCATTTGCGACTTTTGCATGATGACAGTCT 2169

QY 355 GAAAGTGAACATATTCATATGTCGCAAGCTCAATGATCGCCAGGAAACGCGGCA 414  
 DB 2168 TACACAGAAAGTTGGAACAAGTTCAAGCCCTTATGATTTAGTCTAAAGTCTGTGCA 2109

QY 415 GTCACTTGGCAGATGCGCTTGCATGGGTGCTGCGGAGCAAGGAGAGTACGCGCGGA 474  
 DB 2108 ATCTCTAGCTCAATTGGCAATAAGCTTGGGTTTACG---GGAAGAGCTGCGCAAGTTCA 2052

QY 475 TACGCTGACCACTGCATGATTGGTGTCTGCTTCAGTGTGAGCAGCTGGACACACGCTTGA 534

```

Db      2051 AGCTGTTACATCAGCTTTGATTTGGGCAAGTCGTCGCCAACAAATCATTTGAAAATTTGCG 1992
QY      535 TTCACTCACAACACTTGGAGTTTCTGACGCCGAGTT 570
Db      1991 AGCTTTAGAACGTCTGGAATTTACTGACCAAGAATT 1956

RESULT 3
LOCUS   AF075981
DEFINITION
Anopheles gambiae GSS 860 bp DNA linear GSS 29-AUG-2000
typhimurium LT2, Lambda DASH II Salmonella
typhimurium genomic clone 390-T3, genomic survey sequence.
ACCESSION
VERSION AF075981
KEYWORDS
SOURCE GSS.
ORGANISM
Salmonella typhimurium
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
REFERENCE
1 (bases 1 to 860)
Wong, R.M.-Y., Wong, K.K., Benson, N.R. and McClelland, M.
Sample sequencing of a Salmonella typhimurium LT2 lambda library:
comparison to the Escherichia coli K12 genome
FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
9243757
10227170
Contact: McClelland M
Molecular Biology
Sidney Kimmel Cancer Center
3099 Science Park Road, San Diego, CA 92121, USA
Email: mclelland@lifsci.sdsu.edu
Class: shotgun.
FEATURES
source
1..860
Location/Qualifiers
/organism="Salmonella typhimurium"
/mol_type="genomic DNA"
/strain="LT2"
/db_xref="taxon:602"
/clone="390-T3"
/clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
/note="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer"

ORIGIN
Query Match 3.0%; Score 72; DB 28; Length 860;
Best Local Similarity 54.7%; Pred. No. 9.9e-08;
Matches 220; Conservative 0; Mismatches 160; Indels 22; Gaps 3;

QY 52 TGAAGAAGGCTTTTACGTGGTATTTCTTCTACGGTCCAGAGCTCACAGCGAGGCGGC 111
Db 18 TGGCAAGCGTTGTACGTGGGATCTCTAACTATCTGCGGATCTGGCCAGACAGCTAT 77

QY 112 TGAGTTATGCGGAGAGGCGTGGCGCTTCTGATTCATCAGCCAGCTATTCATCAT 171
Db 78 CGATATCTGGAGGATCTGGACGCGCTTGCTGATTCATCAGCCGCTTATTCGCTTTT 137

QY 172 TAATCTTGGGTGGAGAACCGGGCGATGACGGGTGAGAACTTTGTCAGTTCAGTCCCAA 231
Db 138 TGAGCGTTGGGTAGAGAGCGGGC-----TGCTGGGCTGTGTGAGA 179

QY 232 CAATGCTCTTGGCGTCAATTCCTTCTCAACATTGCGCGAGGCGCTGCTCAACGACAAATA 291
Db 180 AAAAAGCGTCGTAGTATTCCTTCTGCGCTGGGGCGGGCA-GCTCACTGACCGTAA 238

QY 292 TCTCGATGTAATTCAGAGGTTCCGCG---GCCAGCGAGGTTAAGTCCCTGTCGAGGG 348
Db 239 TTTGAATGGTATTCGGAAGATTCCCGCGGCGGAGCGGAACCGGTTTCTTAAACACAGA 298

QY 349 CATGTTGAACGTGAACAATATGATATGCTCCGCAAGCTCAATGACATCCGCCAGACG 408
Db 299 ACAGATTACCGCGCAAACTGAAAAGTTGCGCGGTTGCAATGACTGCTCGCGAGCG 358

```

```

QY 409 CGGCGAGTCACCTGGCAGATGGCGCTTGCATGGGTGCTGCG 450
Db 359 GGGTCAAAATTTGATCAATGGCGCTCGCTGATGCTGCG 400

RESULT 4
LOCUS   CNS01MOH
DEFINITION
Anopheles gambiae GSS 830 bp DNA linear GSS 14-JUN-2001
from strain PRST of Anopheles gambiae (African malaria mosquito),
genomic survey sequence.
ACCESSION
VERSION ALI51258
KEYWORDS
SOURCE GSS.
ORGANISM
Anopheles gambiae (African malaria mosquito)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoides;
Anopheles.
REFERENCE
1 (bases 1 to 830)
Roth, C.W., Brey, P.T., Ke, Z., Collins, F.H. and Weissenbach, J.
Genoscope.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 830)
Roth, C.W., Brey, P.T., Ke, Z., Collins, F.H. and Weissenbach, J.
Direct Submission
Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the
Laboratory of Biochem. and Biol. Molec. of Insects, Institut
Pasteur.
FEATURES
source
1..830
Location/Qualifiers
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PRST"
/db_xref="taxon:7165"
/clone="22E24"
/clone_lib="NotreDamel"
/note="end : T7"

ORIGIN
Query Match 2.5%; Score 60.4; DB 29; Length 830;
Best Local Similarity 50.7%; Pred. No. 0.00013;
Matches 142; Conservative 1; Mismatches 137; Indels 0; Gaps 0;

QY 1058 GGGCGCTTCCACCTTTGTTGTTCATGGCGTCTTTTCGTCGCCATGACGGCAACATAACAGG 1117
Db 318 GCCAGTTTCAGGTTACTGCTCATCGCGGCTTTTCAGCGCGCGAATCCGTACCAGAGCAGA 377

QY 1118 TAAGCATGCCACCCAGCGCATATATCGAGCAGCATCGCGCGCATTCGACAAAGA 1177
Db 378 AACGCCACCGCCGCCCGCGGTCGCCAAGCCAGCAGCAGCAGCATCGACGCGCG 437

QY 1178 TCACGCCCCAGGTGCGCGGATGAACAAAAGACGTCAGAAATTAACACACAGAGA 1237
Db 438 CTGCCGCCAAAATCCCGCAGATCAGCAGCAATCACTTACCGCGCATAGCAGGCA 497

QY 1238 ACCGCAATGAGTCTTCGCGCTTAATTCCTTTTAAATCACCAGTACATTCGCGTCCG 1297
Db 498 ATCATCAGATGCTACTGCGCGGATGCGCTGATTCATCAGCAGCATTTTTCGCGCGG 557

QY 1298 ATGGAAGTAAAGACTGGCCCCCAAGAGCAGCTGTA 1337
Db 558 AGGGAAGAAATAAAAAATGCACCTAAGGCAAGCCCTTGAM 597

RESULT 5
LOCUS   BZ554294/c
DEFINITION
Anopheles gambiae GSS 1034 bp DNA linear GSS 17-DEC-2002

```

```

DEFINITION   pacsl-60_4617.x1 pacsl-60 Pseudomonas aeruginosa genomic clone
ACCESSION   BZ554294
VERSION     BZ554294.1 GI:27161466
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
ORGANISM    Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 1034)
AUTHORS     Burns,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE       Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
JOURNAL     J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            University of Washington
            Genome Center
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES     Location/Qualifiers
             1..1034
                /organism="Pseudomonas aeruginosa"
                /mol_type="genomic DNA"
                /strain="1-60"
                /db_xref="taxon:287"
                /clone="pacsl-60_4617"
                /clone_lib="pacsl-60"
                /notes="clinical isolate 1-60 Whole genomic shotgun
                library."
ORIGIN
Query Match      2.4%; Score 56.8; DB 28; Length 1034;
Best Local Similarity 51.2%; Pred. No. 0.0014;
Matches 133; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
QY 1094 GCGATGACGCGCAACCAATACAGGTAGCGTGCACCCCAAGGTGCGCGCGATGACAAAGAACG 1213
DB 341 GCGTTGAGCGCGTACCGGTGAGGAGGCGATTCGCCCGGCGGCGGCGGCGGCGGCGGCGGCG 282
QY 1154 ATCGCGCGCGCATTTGACAAAGATCAACGCCCAAGGTGCGCGCGATGACAAAGAACG 1213
DB 281 GTTGGGTTTTCAGCAGCAGCTTGCGCCAGCCGACAGCGAGGCTGACCAAGACCGCG 222
QY 1214 TCAGAAATTAACACGAGAGAACCGCATAGTCTTCGCGCTTAATTCCTTTGTTA 1273
DB 221 TCGCAGAACACGACGACGCGCGCTACCGAGAGGTGATGCTCGCGCGCGAGGCTCTCGCG 162
QY 1274 ATCCAGTACATTCTGCGGTCGATGACAGTAAAGACTGGCCCGCCCAAGACGACCT 1333
DB 161 AGGACGAGGCGATTTGCGCGCGATGCGATGATCAGCGCGCGCGCGCGCGCGGATCGG 102
QY 1334 GTAATGAGATTTCCATGAT 1353
DB 101 TTGATAGTCTGGAACAT 82

RESULT 6
BZ560550
LOCUS       BZ560550
DEFINITION   pacsl-60_4617.x1 pacsl-60 Pseudomonas aeruginosa genomic clone
ACCESSION   BZ560550
VERSION     BZ560550.1 GI:27178749
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
ORGANISM    Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 897)

```

```

AUTHORS     Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE       Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
JOURNAL     J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            University of Washington
            Genome Center
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES     Location/Qualifiers
             1..897
                /organism="Pseudomonas aeruginosa"
                /mol_type="genomic DNA"
                /strain="2-164"
                /db_xref="taxon:287"
                /clone="pacs2-164_2569"
                /clone_lib="pacs2-164"
                /note="clinical isolate 2-164 Whole genomic shotgun
                library."
ORIGIN
Query Match      2.3%; Score 54.2; DB 28; Length 897;
Best Local Similarity 52.8%; Pred. No. 0.0067;
Matches 143; Conservative 0; Mismatches 120; Indels 8; Gaps 1;
QY 650 THACCATCAACATCAGTTTATGCGCAATGCGGTATCACTCACTGCCAGCAGCTTG 709
DB 290 TATCCCGGAGAACGACGATTCGCGCGCCATGCCAGCATCATGGCGCCACCATCAGGTGCG 349
QY 710 ATCCAGCGCACACCTTGGGGCTGACAGCGCGGTGCAATGCTGTGCGCGCAACCC 769
DB 350 AGCAGCGCCAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 409
QY 770 ACCAGCGGAGAACCATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 829
DB 410 AGCGCGAGAACCATCATCAGCAGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 465
QY 830 GTGTGCGCGTATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 889
DB 466 ----CGCGGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521
QY 890 GGGTTCAACAGGTTCAGCAGCATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 920
DB 522 GGGTTAAGCCAGGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 552

RESULT 7
BZ568946
LOCUS       BZ568946
DEFINITION   pacs2-164_8165.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
ACCESSION   BZ568946
VERSION     BZ568946.1 GI:27202770
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
ORGANISM    Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 1620)
AUTHORS     Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE       Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
JOURNAL     J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            University of Washington
            Genome Center
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954

```



ORIGIN		2.0%;	Score 48.6;	DB 29;	Length 925;
Query Match		13.8%;	Pred. No. 0.22;		
Best Local Similarity		Matches 49;	Conservative 157;	Mismatches 149;	Indels 0;
		Gaps 0;			
QY	681	CGGTTCATCAACTGCCACGACGAGCGTTTGATCCAGCGCCACACCTTGGGGCTGGACAGCG	740		
Db	925	CSBSCSCSCSBSBSCSSSMSTSSNSBSCSSBSSSSTSSMSSSBSSBSSSGSSSS	866		
QY	741	GGCTGACATCTCTGCTGGCCGAAACCCACACGCGGAAACGATACGGCTTGCCCGCA	800		
Db	865	SSGTSTACVKNASSSSCGCCGCMACCMGSCSSSSSCGASAGRVKVRASGAGAKRGGS	806		

[illegible]

Qy	866	ACAAACGGCTCCAAATA	CGCATTCGGGTTCAAC	AGGTCAGCACGATTGCCATCAACATG	925
Db	187	TTCGACGAGGCCACGGG	CGCCTTCGAGATCCACT	CGCGGCACCTGCTACGTCCACTTC	246
Qy	926	GGCTTTTACCAAACCGC	TGCTTTATCGAGCTCACT	CCACCGGCACCGGTTGCGGTG	985
Db	247	GGCTCCCACTCGTCTCT	ACGAGAGGACCATAACGG	CAAGCTCTCCAAAGGCGGCATC	306
Qy	986	TCAGTGGCCACCGCCGA	ACCGGCCAA	1011	
Db	307	TCGGACCTCTCCGGCGT	CTCAGGCCAA	332	

RESULT 13				
CD881613/c				
LOCUS	CD881613	720 bp	mRNA	linear EST 14-JUL-2003
DEFINITION	F1.103123F010329 F1	Triticum aestivum	cDNA clone F1103123,	mRNA sequence.

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Pooideae; Triticeae; Triticum.

COMMENTS

Contact: Genoplatte  
Genoplatte  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french  
plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>  
and <http://genoplatte-info.inbio.gen.fr>).

```

/cultivar="recital"
/db_xref="taxon:4565"
/_clone="F103L23"
/tissue_type="leaf one"
/_clone_lib="F1"

```

		520	ATGGCCAGCTCGGCTCCCAAGCTTTTTCAGGACTCCAGCTTGGGGTGTCTGGCCGTGGCC	461
	Db			
		746	GACATGCTGCTCGCGCCGAACCCACGACGGGGAAACAGATCAGGCTTCCCGCGGAACGCC	805
	Qy			
		460	GCCACCGCTCGACGCGCGGTAACTCTTTCGCGACGCTGGATGACAGAGGGTCCCGACTCGC	401
	Db			

[illegible][illegible]

RESULT 14  
 C829022  
 LOCUS 620 bp mRNA linear EST 11-DEC-2002  
 DEFINITION 1114036D07.y2 1114 - Unigene IV from Maize Genome Project Zea mays  
 ACCESSION C829022  
 VERSION C829022  
 KEYWORDS Zea mays  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 Zea mays  
 Walbot V.  
 Maize ESTs from various cDNA libraries sequenced at Stanford  
 University  
 Unpublished (1999)  
 CONTACT: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 1114036 row: D column: 07.

FEATURES  
 source  
 1..620  
 /location/Qualifiers  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4577"  
 /clone\_lib="1114 - Unigene IV from Maize Genome Project"  
 /note="This library represents the unique genes found in  
 the fourth round of EST sequencing at Stanford University  
 for the maize genome project. Sequences are present from  
 libraries 1091 and 3524. Contigs were assembled using  
 ZmDBAssembler and 2 representatives from each contig were  
 selected for the Unigene set. All singlets were also  
 selected."

ORIGIN  
 Query Match 1.9%; Score 44.4; DB 14; Length 620;  
 Best Local Similarity 46.0%; Pred. No. 2.5;  
 Matches 150; Conservative 0; Mismatches 176; Indels 0; Gaps 0;  
 QY 686 ATCAACTGCTCCAGACGAGTGTATCCAGCGCCACACTTGGGGCTGGACAGCGGGGT 745  
 DB 32 ATCCCAATATCCGACAGCGCAATGTCTCCGCTCCCGGATCGTCTCTCGCGTCC 91  
 QY 746 GAACAATGCTGTGGCGGAAACCCAGCGGGAACAGATCAGGCTTGGCGGCAACGCG 805  
 DB 92 GCGCGCGCGCGCGAGCGGTACAGCGGTACAGCGGTACAGCGGTACAGCGGTACAG 151  
 QY 806 CCAGCGGCGAAATCCACGCTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGT 865  
 DB 152 CTACTCCCAAGTACGCGCTCCCGAAGGGGCTCATCCCGGACTCCGCTCTCTACAGC 211  
 QY 866 ACAACAGCGCTCCAAATACGATTCGGGTTCACACGAGTCAGCAGGATTCGATCAACATG 925  
 DB 212 TTGAGAGGCGCAGCGGCGCTTGGAGTTCACCTCCGCGGACCTCTGCTACGTCACATTC 271  
 QY 926 GCGTTTACCCAAACCCGCTGTATTCGAGCTCACTCCACCCGACCCGCTGTGGCGGTG 985  
 DB 272 GGTCTCCACCTCTGCTACTACGAGGAGACATAACCGGACGCTCTCCAGGGGCGCCATC 331  
 QY 986 TCAGTGGCCACCGCGGACCGGCA 1011  
 DB 332 TCGGACCTCTCCGCGGTCCAGGCCAA 357

RESULT 15  
 BQ704105

LOCUS BQ704105  
 DEFINITION 946110C04.y1 946 - tassell primordium prepared by Schmidt lab Zea  
 mays cDNA, mRNA sequence.  
 ACCESSION BQ704105  
 VERSION BQ704105.1 GI:21843524  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 Zea mays  
 Walbot V.  
 Maize ESTs from various cDNA libraries sequenced at Stanford  
 University  
 Unpublished (1999)  
 CONTACT: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 946110 row: C column: 04.  
 Location/Qualifiers  
 1..644  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="OH43"  
 /db\_xref="taxon:4577"  
 /tissue\_type="tassels"  
 /dev\_stage="just after the transition from vegetative to  
 inflorescence development"  
 /lab\_host="XLOR"  
 /clone\_lib="946 - tassell primordium prepared by Schmidt  
 lab"

ORIGIN  
 Query Match 1.9%; Score 44.4; DB 13; Length 644;  
 Best Local Similarity 46.0%; Pred. No. 2.5;  
 Matches 150; Conservative 0; Mismatches 176; Indels 0; Gaps 0;  
 QY 686 ATCAACTGCTCCAGACGAGTGTATCCAGCGCCACACTTGGGGCTGGACAGCGGGGT 745  
 DB 39 ATCCCAATATCCGACAGCGCAATGTCTCCGCTCCCGGATCGTCTCTCGCGTCCG 98  
 QY 746 GAACAATGCTGTGGCGGAAACCCAGCGGGAACAGATCAGGCTTGGCGGCAACGCG 805  
 DB 99 GCGCGCGCGCGCGAGCGGTACAGCGGTACAGCGGTACAGCGGTACAGCGGTACAG 158  
 QY 806 CCAGCGGCGAAATTCACGCTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGT 865  
 DB 159 CTACTCCCAAGTACGCGCTCCCGAAGGGGCTCATCCCGGACTCCGCTCTCTACAGC 218  
 QY 866 ACAACAGCGCTCCAAATACGATTCGGGTTCACACGAGTCAGCAGGATTCGATCAACATG 925  
 DB 219 TTGAGAGGCGCAGCGGCGCTTGGAGTTCACCTCCGCGGACCTCTGCTACGTCACATTC 278  
 QY 926 GCGTTTACCCAAACCCGCTGTATTCGAGCTCACTCCACCCGACCCGCTGTGGCGGTG 985  
 DB 279 GGTCTCCACCTCTGCTACTACGAGGAGACATAACCGGACGCTCTCCAGGGGCGCCATC 338  
 QY 986 TCAGTGGCCACCGCGGACCGGCA 1011  
 DB 339 TCGGACCTCTCCGCGGTCCAGGCCAA 364

Search completed: March 16, 2004, 02:52:19



Job time : 6529.5 secs